

Quantum channels and memory effects

Filippo Caruso*

*Institut für Theoretische Physik, Universität Ulm, Albert-Einstein-Allee 11, D-89069 Ulm, Germany, and
LENS and Dipartimento di Fisica e Astronomia, Università di Firenze, 50019 Sesto Fiorentino, Italy*

Vittorio Giovannetti†

NEST, Scuola Normale Superiore & Istituto Nanoscienze-CNR, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

Cosmo Lupo‡ and Stefano Mancini§

School of Science and Technology, University of Camerino, I-62032 Camerino, Italy

Any physical process can be represented as a quantum channel mapping an initial state into a final state. Hence it can be characterized from the point of view of communication theory. Recently a lot of efforts have been devoted to encompass memory effects within this approach. It is usually meant that memory effects arise when the application of a channel on several inputs is not merely the result of identical and independent maps. Alternatively, if the channel is defined by its infinitesimal generator in a continuous-time description, one refers as well to memory effects if such an evolution is non-Markovian. The consideration of such effects has given rise to more powerful tools than the traditional quantum channel approaches for describing quantum information processes. Moreover it constitutes a fertile ground where interesting novel phenomena emerge at the intersection of quantum information theory with other branches of physics. Recent developments in the field of quantum channels and memory effects are here reviewed.

PACS numbers: 03.67.-a, 89.70.-a, 02.50.-r

Contents

I. Introduction	2	1. Performance achieved in the case of correlated errors	13
II. Quantum channels: basic properties	3	B. Decoherence free subspaces	15
A. Degradable, PPT, and Entanglement Breaking channels	4	1. Threshold values for the degree of correlation	16
B. Causal, semi-causal, and localizable channels	4	C. Designing quantum codes for correlated errors	16
III. From Memoryless to Memory Quantum Channels	5	1. Concatenated codes	16
A. Memoryless quantum channels	5	2. Burst errors codes	17
B. Non-anticipatory memory quantum channels	5	3. Convolutional codes	18
1. Structure theorems for non-anticipatory quantum channels	6	VI. Coding Theorems and Capacities	19
2. Finite-memory channels	7	A. Operational definitions and memoryless setting	19
3. Symbol independent vs intersymbol interference maps	7	B. Entropic upper bounds: memory setting	20
C. Quasi-local algebras approach	8	1. Finite-memory channels	21
1. The Structure of Causal Channels	8	2. Perfect memory channels	21
2. Ergodic channels with decaying input memory	8	3. Non-anticipatory memory channels	21
IV. Taxonomy of Quantum Memory Channels	8	4. Forgetful channels	22
A. Localizable, fully non-anticipatory memory quantum channels	9	5. Long-term memory channels	22
B. Perfect memory channels	9	6. Ergodic cq-channels with decaying input memory	23
C. Markovian channels	10	VII. Solvable Models of Memory Channels	23
D. Fixed-point, indecomposable, and forgetful channels	11	A. Discrete memory channels	23
E. Long term memory channels	12	B. Continuous memory channels	26
V. Quantum Codes	12	VIII. Quantum channels from dynamics with memory	30
A. Standard quantum coding theory	12	A. Non Markovian master equations	31
		B. Legitimate memory kernels	32
		C. Markovian vs non-Markovian dynamics	32
		D. A solvable model	33
		IX. Summary and Outlook	33
		Acknowledgments	34
		A. Tools for characterizing quantum channels	34
		B. Entropic quantities	35
		C. Decomposition for non-anticipatory quantum channels	36

*Electronic address: filippo.caruso@lens.unifi.it

†Electronic address: v.giovannetti@sns.it

‡Electronic address: cosmo.lupo@unicam.it

§Electronic address: stefano.mancini@unicam.it

I. INTRODUCTION

Any physical process implies a state change. Hence the “channel” formalism, commonly used in communication theory to map input to output information, is suitable for describing also physical processes, provided the input (initial) and output (final) states are regarded in terms of their information content. As consequence physical processes, viewed as channels, can be characterized by their information transmission ability. The seminal work of Shannon (Shannon, 1948; Verdú, 1998) showed that a memoryless noisy channel, where the noise acts independently on each symbol sent through the channel, can be parameterized by a single quantity, the *capacity* of the channel. Shannon defined the capacity as the maximum rate at which information can be reliably sent through a channel, with the use of suitable coding, and proved its expression in terms of entropic functions of the input and output systems. The work of Shannon has been extended to include channels with “memory”, where the noise is no longer independent and identical on different inputs (Gallager, 1968).

Nevertheless, quantum information theory has taught that it is the laws of physics that establish the ultimate limits in processing and communicating information (Nielsen and Chuang, 2000; Petz, 2008). That is why quantum mechanics, the most advanced physical theory, comes into play. For quantum channels, several notions of capacities have been introduced, depending on whether classical or quantum information is conveyed and if additional resources, such as pre-shared entanglement, are present. A straightforward extension of the Shannon results leads to the so-called entanglement assisted classical capacity (Adami and Cerf, 1997; Bennett *et al.*, 1999), i.e., the capacity in transmitting classical symbols encoded into quantum states assisted by pre-shared entanglement between the sender and the receiver (Bennett and Shor, 1998). The classical capacity of quantum channels without any side resource has been proved to be expressed in terms of the accessible classical information (Holevo, 1998; Schumacher and Westmoreland, 1997). Additionally the quantum capacity, i.e., the capacity in transmitting quantum resources as coherence and entanglement, has been considered (Devetak, 2005; Lloyd, 1997).

While the majority of the work in quantum channels has been initially concerned with the study of maps that do not exhibit memory effects, the assumption of vanishing correlations — or memory — in the system-environment interaction is not always justified. For example, with increasing transmission rates in communication channels, successive transmissions happen so rapidly that the environment may retain a memory between individual transmissions. Optical fibers are an example in

which such effects may be explored experimentally (Ball, Dragan and Banaszek, 2004; Banaszek *et al.*, 2004). In quantum information processors, especially in solid state implementations, qubits may be so closely spaced that individual elements of the environment will interact with several qubits, thus leading to cross-talks and correlations in the noise (Hu, 2007). The consideration of spatial and temporal memory effects is therefore becoming increasingly pressing with the continuing miniaturization of information processing devices and with increasing communication rates through channels.

In recent years an increasing attention has been devoted to quantum channels with memory. Apparently, the interest towards information transmission through quantum channels with memory spread after the model introduced by (Macchiavello and Palma, 2002). Macchiavello and Palma, 2002 gave an example of a qubit channel with Markovian correlated noise in which the encoding of information by means of entangled input states may increase the transmission rate of classical information. Such an increment in the communication rate due to entanglement is strictly related to the property of superadditivity of quantum channels for the transmission of classical information (Hastings, 2009). Subsequently, the study of quantum channels with memory has largely been confined to channels with Markovian correlated noise with the aim of deriving bounds on the classical capacity (see e.g. (Bowen and Mancini, 2004; Bowen, Devetak and Mancini, 2005; Hamada, 2002)). Then, coding theorems have been devised for wide classes of quantum memory channels having structural properties that guarantee “regular” asymptotic behavior (Bjelaković and Boche, 2008; Datta and Dorlas, 2007; Kretschmann and Werner, 2005). This fact can be traced back to the work of (Dobrushin, 1963), who showed that a limiting expression of maximal normalized mutual information holds for a wide class of channels exhibiting stationary or ergodic behavior. A completely different approach was taken by Hayashi and Nagaoka (Hayashi and Nagaoka, 2003), that applied the “information-spectrum” method to obtain a coding theorem for the classical capacity, following the work by Verdú and Han, 1994 on classical channels with memory.

A lot of work has been undertaken also along the direction of bosonic Gaussian memory channels. Inspired by the seminal paper of Shannon (Shannon, 1949) on classical memory channels generalizing the formula for the capacity of (power-constrained) white Gaussian channels to dispersive/nonwhite Gaussian channels by the “water-filling” formula, quantum “water-filling” formulae have been recently derived (Pilyavets, Lupo and Mancini, 2009; Schäfer, Karpov and Cerf, 2009). Moreover, rigorous justification of the water-filling formula for classical dispersive/nonwhite Gaussian channels usually appeals to the Toeplitz distribution theorem (Gray, 1972). Then, the latter has been used in (Lupo, Giovannetti and Mancini, 2010) for evaluating capacities of a general bosonic Gaussian memory channels that can be “unrav-

elled” into memoryless ones by normal mode decomposition.

Aside from that, models of memory channels were introduced in which memory effects arise from the interaction with an environment in a correlated state (Giovannetti and Mancini, 2005), allowing one remarkable links between capacity of quantum channel and many-body physical properties of the environment, like phase transitions (Plenio and Virmani, 2007, 2008).

Moving towards the dynamics of open quantum systems it is worth noting a Lindbladian approach to memory channels taken in (Daffer, Wódkiewicz and McIver, 2003; Daffer *et al.*, 2004). Then in the continuous-time description of open quantum systems considerable attention has been devoted to the study of quantum channels arising from non-Markovian dynamics. A different terminology is commonly used in this framework, since Markovian-correlated dynamics correspond to the memoryless setting, and dynamics with memory are associated with non-Markovian effects. A typical approach to the dynamics of open quantum systems uses the Nakajima-Zwanzig projection operator technique (Breuer and Petruccione, 2002; Nakajima, 1958; Zwanzig, 1960) which shows that under fairly general conditions, the master equation for the reduced density operator takes the form of a nonlocal equation in which memory effects are taken into account through the introduction of a memory kernel. Then, the problem is to find those conditions on the memory kernel ensuring that the time evolution map is a *bona fide* quantum channel (Chruściński and Kossakowski, 2012). Contrarily to the Markovian case, a full characterization of legitimate memory kernels is still missing. Moreover, the concept of non Markovianity in this framework is itself not uniquely defined (Breuer, Laine and Piilo, 2009; Rivas, Huelga and Plenio, 2010).

The paper is organized as follows: The notation and the basics properties of quantum channels are presented in Sec. II. Section III formalizes the notion of memory quantum channels and discusses the physical mechanisms that are responsible for the arising of correlations in quantum communication. The various classes of memory channels are reviewed in Sec. IV. Section V discusses quantum error correction codes and achievable information transmission rates. Coding theorems and capacities are presented in Sec. VI. Section VII concerns models of memory channels that can be solved for their communication capacities. Then, quantum channels arising from dynamics with memory are discussed in Sec. VIII. Finally, conclusion and outlook are given in Sec. IX.

II. QUANTUM CHANNELS: BASIC PROPERTIES

In a typical communication scenario two parties (Alice the sender of the message and Bob the receiver) try to exchange (classical or quantum) information by encoding it into (possibly infinitely long) sequences of signals which

propagate through the medium that separate them. A train of transmitted signals defines a sequence of independent uses of the communication line (*channel uses*), and their input-output evolution from Alice to Bob is determined by the noise which tampers with the transmission process. In quantum information the channel uses are represented by the degree of freedom (e.g., polarization, spins) of a collection $Q := \{q_1, q_2, \dots\}$ of identical information carrying objects (e.g., optical pulses, flying atoms or ions) which are locally produced by Alice and organized in a time-ordered sequence. In this setting the noise can then be described by assigning a proper *mapping* which acts on the (global) input states of the information carriers Q to produce the associated (global) output states received by Bob, i.e.

$$\rho_Q \in \mathfrak{S}(\mathcal{H}_Q) \mapsto \rho_{Q'} = \Phi(\rho_Q) \in \mathfrak{S}(\mathcal{H}_{Q'}), \quad (1)$$

[here $\mathfrak{S}(\mathcal{H}_Q)$ and $\mathfrak{S}(\mathcal{H}_{Q'})$ stand for the sets of the density matrices associated with the Hilbert spaces $\mathcal{H}_Q, \mathcal{H}_{Q'}$ which describe the carries as seen by Alice and Bob respectively – the two need not to coincide in general]. Secs. III and IV will show in which way memory effects can arise in these processes focusing on the features of the transformation Φ that are responsible for introducing correlations among the various components of Q . Before doing so, however, it is useful to remind that quantum mechanics poses some fundamental structural constraints on any transformations of the form (1), which must apply independently from the underlying physical mechanisms that is governing the process and independently from the composite nature of input system Q . In particular, assuming that prior to the communication, no correlations exist between Q and the channel medium, the mapping Φ must be

- i) *linear* [i.e. it must transform mixtures of input density matrices into the corresponding mixture of the associated outputs, i.e. $\sum_i p_i \rho_Q(i) \mapsto \sum_i p_i \Phi(\rho_Q(i))$, with p_i the probability associated with the input state $\rho_Q(i)$. This in particular allows one to expand Φ as a linear super-operator on the full algebra $\mathcal{B}(\mathcal{H}_Q)$ of the bounded operators of \mathcal{H}_Q];
- ii) *completely positive* [i.e. when acting on Q it must preserve the positivity of any initial joint density matrix ρ_{QA} of Q with an arbitrary external ancilla A that is NOT propagating along the channel];
- iii) *trace-preserving*, [i.e. it must preserve the normalization of all input states];

a violation of any of these three conditions implying the impossibility of maintaining the statistical interpretation of the theory (Breuer and Petruccione, 2002; Holevo, 2011; Petz, 2008). In the jargon of quantum information theory, any transformations (1) that fulfills the three conditions given above is said CPTP (completely positive and trace preserving) mapping or simply *quantum*

channel. While referring the reader to (Bengtsson and Życzkowski, 2006; Breuer and Petruccione, 2002; Holevo, 2011; Holevo and Giovannetti, 2012; Keyl, 2002; Petz, 2008) for an exhaustive characterization of their properties, one reminds that CPTP maps form a closed set under convex combination and a semigroup under super-operator composition, [of course when proper correspondences between the associated input and output spaces applies]. It is also well known that any quantum channel Φ typically admits unitary dilations that allow one to represent it in terms of a unitary coupling with an external (possibly fictitious) environment E which not necessarily corresponds to the channel medium. For instance, taking for simplicity $\mathcal{H}_Q = \mathcal{H}_{Q'}$, one can write

$$\Phi(\rho_Q) = \text{Tr}_E \left[U_{QE} (\rho_Q \otimes |\omega\rangle_E \langle \omega|) U_{QE}^\dagger \right], \quad (2)$$

where $|\omega\rangle_E$ is a fix state of E , U_{QE} is the unitary transformation which couples the latter to the input system Q , and Tr_E denotes the partial trace over environment. Equation (2) can also be put in correspondence with the *operator sum* (or *Kraus*) representation of Φ ,

$$\Phi(\rho_Q) = \sum_{j=1}^{d_E} K_j \rho_Q K_j^\dagger, \quad (3)$$

by identifying d_E with the dimension of E and the Kraus operator K_j with the linear operator ${}_E \langle j | U_{QE} | \omega \rangle_E$ of \mathcal{H}_Q , where $\{|j\rangle_E; j = 1, \dots, d_E\}$ is an orthonormal basis of E . There exists finally a fundamental relation, known as the Choi-Jamiolkowski (CJ) isomorphism (Choi, 1975; Jamiolkowski, 1972), which permits to faithfully represent any CPTP Φ as a density matrix of a composite system QA with A being an auxiliary system isomorphic to Q . The explicitly connection is obtained by applying the map Φ to half of a maximally entangled state $|\beta\rangle_{QA}$ of QA to create the so called CJ state of the channel,

$$\rho_{QA}^{(\Phi)} := (\Phi \otimes \text{id})(|\beta\rangle_{QA} \langle \beta|), \quad (4)$$

[here id stands for the identity map on A].

Equation (1) implicitly assumes the Schrödinger picture in which the states of the system are evolved while the observable are kept fixed. In the Heisenberg picture, in which instead the states are fixed and the observables evolve in time, the channel is represented by the dual map $O \mapsto \Phi^*(O)$ defined on the space of the receiver observable O and identified through the identity

$$\text{Tr}[\Phi(\rho_Q)O] = \text{Tr}[\rho_Q \Phi^*(O)]. \quad (5)$$

As its Schrödinger counterpart, the transformation Φ^* is linear and completely positive, but in general it is not trace preserving. On the other hand it is always *unital*, i.e., it maps the identity operator into itself. Operator sum representations for Φ^* can be easily constructed from those of Φ (3), yielding,

$$\Phi^*(O) = \sum_{j=1}^{d_E} K_j^\dagger O K_j. \quad (6)$$

A. Degradable, PPT, and Entanglement Breaking channels

Associated with the unitary dilation (2) is the notion of the *complementary channel* of Φ . The latter is the CPTP mapping $\tilde{\Phi} : \mathfrak{S}(\mathcal{H}_Q) \rightarrow \mathfrak{S}(\mathcal{H}_E)$ which moves the initial states of the system Q into the states of the environment E via the transformation

$$\rho_Q \mapsto \tilde{\Phi}(\rho_Q) := \text{Tr}_Q \left[U_{QE} (\rho_Q \otimes |\omega\rangle_E \langle \omega|) U_{QE}^\dagger \right], \quad (7)$$

where now Tr_Q denotes the partial trace over the system Hilbert space. This definition allows us to introduce another property of quantum channels, which is called *degradability* (Devetak and Shor, 2005). Loosely speaking, a map is *degradable* when one can recover the final environment state $\tilde{\Phi}(\rho_Q)$ just applying a third CPTP map to the output system state. More rigorously, a *degradable* map is such that there does exist \mathcal{T} CPTP satisfying the relation:

$$\tilde{\Phi}[\rho_Q] = (\mathcal{T} \circ \Phi)(\rho_Q) := \mathcal{T}(\Phi(\rho_Q)), \quad (8)$$

for any input state ρ_Q [here “ \circ ” stands for the super-operator composition]. Similarly, a channel is called *anti-degradable* when the opposite relation holds, i.e.,

$$\Phi[\rho_Q] = (\mathcal{T} \circ \tilde{\Phi})(\rho_Q) := \mathcal{T}(\tilde{\Phi}(\rho_Q)). \quad (9)$$

Another important class of quantum channels is represented by the positive partial transpose (PPT) or binding maps. These are CPTP transformations $\Phi : \mathfrak{S}(\mathcal{H}_Q) \rightarrow \mathfrak{S}(\mathcal{H}_Q)$ which, when operating on joint states ρ_{QA} produce outputs having positive partial transpose, or equivalently, which possess a positive partial transpose CJ state (Horodecki *et al.*, 1996; Peres, 1996; Rains, 2001). A proper subset of PPT channels is represented by the so-called entanglement breaking channels, characterized by a separable CJ state (Horodecki, Shor and Ruskai, 2003; Ruskai, 2003).

B. Causal, semi-causal, and localizable channels

A couple of interesting notions that have been developed in Refs. (Beckman *et al.*, 2001; Eggeling, Schlingemann and Werner, 2001; Piani *et al.*, 2006) for the special case in which Q can be treated as a finite bipartite system, e.g. by splitting the set of carriers $\{q_1, q_2, \dots\}$ into two independent subgroups Q_1, Q_2 so that $Q = Q_1 Q_2$. In this context the transformation Φ of Eq. (1) is said to be $Q_1 \rightarrow Q_2$ *semicausal* (Beckman *et al.*, 2001) if for any local quantum channel Ψ applied to Q_1 before the action of the map Φ , there is no detectable effect in the subsystem Q_2 , i.e.

$$\text{Tr}_{Q_1} [\Phi(\rho_Q)] = \text{Tr}_{Q_1} \{ \Phi[(\Psi \otimes \text{id})(\rho_Q)] \}, \quad (10)$$

where ρ_Q is a generic [possible entangled] input state of the two carriers and where $\text{Tr}_{Q_1}[\dots]$ describes the partial trace with respect to Q_1 . In other words, for

$Q_1 \nrightarrow Q_2$ *semicausal* maps cross-talking from Q_1 to Q_2 is prevented. Similarly, one may introduce the notion of $Q_2 \nrightarrow Q_1$ *semicausal* map. When both properties are satisfied, the map is called *causal* or *non-signaling*. Special examples of non-signaling channels are the tensor product channels which can be expressed as $\Phi = \Phi_1 \otimes \Phi_2$ with $\Phi_{1,2}$ being CPTP maps operating locally on Q_1 and Q_2 respectively.

Still remaining in the bipartite scenario discussed above Φ is said to be *localizable* if it can be implemented by means of a bipartite state shared between two parties that operates locally on the subsystems Q_1 and Q_2 without communicating (Beckman *et al.*, 2001). That is, localizable channels do not need communication between Q_1 and Q_2 to be simulated. Formally, this is the case when Φ satisfies the following property

$$\Phi(\rho_{Q_1 Q_2}) = \text{Tr}_{A_1 A_2} [(\Psi \otimes \Omega)(\rho_{Q_1 Q_2} \otimes \omega_{A_1 A_2})], \quad (11)$$

where $\omega_{A_1 A_2}$ is a shared bipartite state, and Ψ and Ω are now quantum channels acting locally on subsystems $Q_1 A_1$ and $Q_2 A_2$, respectively. Otherwise, Φ is called $Q_1 \rightarrow Q_2$ *semilocalizable* if one-way communication from Q_1 to Q_2 is required. Accordingly in this case Eq. (11) is replaced by

$$\Phi(\rho_{Q_1 Q_2}) = \text{Tr}_A [(\Psi \circ \Omega)(\rho_{Q_1 Q_2} \otimes \omega_A)], \quad (12)$$

with ω_A being the state of an ancillary system A that acts as the mediator between Q_1 and Q_2 , while Ψ and Ω are quantum channels acting on the systems $Q_2 A$ and $Q_1 A$ respectively. Analogously, one defines a $Q_2 \rightarrow Q_1$ *semilocalizable* map.

By comparison of (11) and (12) it follows that all $Q_1 \rightarrow Q_2$ semilocalizable maps are $Q_2 \nrightarrow Q_1$ semicausal, which in turn implies that all localizable maps are causal. Moreover, it can be proven that semicausality implies semilocalizability, hence semicausal and semilocalizable maps coincide, although causal and localizable maps do not (Beckman *et al.*, 2001; Eggeling, Schlingemann and Werner, 2001; Piani *et al.*, 2006).

III. FROM MEMORYLESS TO MEMORY QUANTUM CHANNELS

Having in mind the multi-uses communication scheme detailed at the beginning of Sec. II in which a time-ordered sequence of carriers $Q := \{q_1, q_2, \dots\}$ propagates from Alice to Bob along a noisy channel, in this section one starts reviewing the mechanisms that are responsible for the introduction of correlations [memory effects] among the various elements of Q .

A. Memoryless quantum channels

Memoryless quantum channels describe those scenarios in which the noise acts *identically* and *independently* on

each element of the sequence Q . Under this assumption the multi-use map associated with the communication line is expressed as a tensor product of a CPTP map $\Phi : \mathfrak{S}(\mathcal{H}_q) \mapsto \mathfrak{S}(\mathcal{H}_q)$ that acts on the states of a single carrier q . Therefore indicating by $\mathcal{H}_Q^{(n)} := \mathcal{H}_q^{\otimes n} = \mathcal{H}_{q_1} \otimes \dots \otimes \mathcal{H}_{q_n}$ the Hilbert space of the first n carriers of the system, its input density matrices $\rho_Q^{(n)} \in \mathfrak{S}(\mathcal{H}_Q^{(n)})$ will be mapped into

$$\Phi^{(n)}(\rho_Q^{(n)}) = \Phi^{\otimes n}(\rho_Q^{(n)}), \quad (13)$$

with $\Phi^{\otimes n} := \Phi \otimes \dots \otimes \Phi$. Equivalently, we can say that the Kraus operators of the memoryless map $\Phi^{(n)}$ can be expressed as a tensor products $K_{i_1} \otimes \dots \otimes K_{i_n}$ formed by independent and identically distributed sequences extracted from the Kraus set $\{K_i\}_i$ associated with the single carrier channel Φ . For an explicit example consider for instance the model discussed in Ref. (Giovannetti, 2005). Here the carriers Q are assumed to propagate from Alice to Bob, one by one and at constant speed, while interacting with an external environmental system via a constant coupling described by the unitary operator $U_{qe} \in \mathcal{B}(\mathcal{H}_q \otimes \mathcal{H}_e)$. The environment e is also assumed to undergo a dissipative process which, on a time-scale τ tends to reset it into a stable configuration ω_e . The memoryless regime is achieved in the limit in which the rate ν at which the carriers propagate from Alice to Bob, is much lower than the inverse of the relaxation time τ , i.e., $\nu \ll 1/\tau$. In this case in fact each carrier couples with identical environmental states. Defining then $\omega_E^{\otimes n} := \omega_{e_1} \otimes \dots \otimes \omega_{e_n}$, this allows us to write

$$\rho_Q^{(n)} \mapsto \text{Tr}_E \left[U_{q_n, e_n} \otimes \dots \otimes U_{q_1, e_1} (\rho_Q^{(n)} \otimes \omega_E^{\otimes n}) U_{q_1, e_1}^\dagger \otimes \dots \otimes U_{q_n, e_n}^\dagger \right], \quad (14)$$

which reduces to Eq. (13) when identifying $\text{Tr}_e[U_{qe}(\dots \otimes \omega_e)U_{qe}^\dagger]$ with unitary dilation of the single-use channel Φ .

B. Non-anticipatory memory quantum channels

Whenever the tensorial decomposition of Eq. (13) doesn't apply, one can speak of *memory channels* or *correlated noise channels*. Among the plethora of possibilities, the following will focus only on those configurations which have physical relevance and which have attracted some interest in the recent literature. In particular, one shall treat those models in which the noise respects the time-ordering of the carriers Q so that at a given channel use, the output cannot be influenced by successive inputs. Under this condition there must exist a family of CPTP maps $\mathcal{F} := \{\Phi^{(n)}; n = 1, 2, \dots\}$ with $\Phi^{(n)} : \mathfrak{S}(\mathcal{H}_Q^{(n)}) \rightarrow \mathfrak{S}(\mathcal{H}_Q^{(n)})$ which allows us to express the output states of the first n carriers in terms of the density matrices of their associated inputs, i.e.,

$$\rho_Q^{(n)} \mapsto \Phi^{(n)}(\rho_Q^{(n)}). \quad (15)$$

Clearly the property (15) requires that the family \mathcal{F} must fulfill the minimal consistency requirement that for all $m < n$ the element $\Phi^{(m)}$ should be obtained as a restriction of $\Phi^{(n)}$ over the degree of freedom of the first m carriers. That is, given $\rho_Q^{(n)} \in \mathfrak{S}(\mathcal{H}_Q^{(n)})$ and $\rho_Q^{(m)} \in \mathfrak{S}(\mathcal{H}_Q^{(m)})$ its reduced density matrix associated with the first m element of Q , one must have

$$\Phi^{(m)}(\rho_Q^{(m)}) = \text{Tr}^{(m)} \left[\Phi^{(n)}(\rho_Q^{(n)}) \right], \quad (16)$$

where $\text{Tr}^{(m)}$ stands for the partial trace over all the carriers but the first m . Borrowing from the classical theory of communication (Gallager, 1968) one can dub these communication lines *non-anticipatory* quantum channels. In Ref. Kretschmann and Werner, 2005 non-anticipatory channels are also called *causal*, it is worth stressing however that this definition of causality has nothing to do with the one introduced in Sec. II.B.

1. Structure theorems for non-anticipatory quantum channels

In the language introduced in Sec. II.B non-anticipatory channels can be classified as *semicausal* with respect to the natural ordering of the channels uses. A direct generalization of the *semicausal/semilocal* connection discussed in Refs. (Beckman *et al.*, 2001; Egging, Schlingemann and Werner, 2001; Kretschmann and Werner, 2005; Piani *et al.*, 2006) can then be used to show that these channels admits a *semilocal* representation where each carrier couples sequentially with a common *memory system* M , whose back-action on the message state simulates the memory effects of the transmission. Accordingly, all the non-anticipatory CPTP maps can be expressed as

$$\Phi^{(n)}(\rho_Q^{(n)}) = \text{Tr}_M \left[U_{q_n M} \cdots U_{q_1 M} (\rho_Q^{(n)} \otimes \Omega_M) U_{q_1 M}^\dagger \cdots U_{q_n M}^\dagger \right], \quad (17)$$

where for all $j = 1, 2, \dots, n$, $U_{q_j M}$ is a unitary transformation which describes the coupling of the j -th carrier with the memory system M , and where Ω_M is some fixed state of M , see Fig. 1 a). An explicit proof of Eq. (17) was first given in Ref. Kretschmann and Werner, 2005 in the context of quasilocal algebras, under the assumption of translational invariance of the noise [more on this will be provided in Sec. III.C]. An alternative proof can also be found in Appendix C

In Eq. (17) M is in general a large system whose dimension d_M is an explicit function of n (in any case it can always be chosen to be less than or equal to d^{2n} with $d = \dim \mathcal{H}_q$ being the dimension of a single carrier). As a matter of fact, as explained in Appendix C one can take M to be a composite system of components m_1, m_2, \dots, m_n whose dimensions can always be chosen to be not larger than d^2 . In this configuration then one can assume Ω_M to be a pure tensor product state of local

terms $|0\rangle_{m_1} \otimes \cdots \otimes |0\rangle_{m_n}$, and write $U_{q_j M}$ as a transformation which couples the j -th carrier *only* with the first j elements of M , i.e.,

$$U_{q_j M} = I_{m_n} \otimes \cdots \otimes I_{m_{j+1}} \otimes U_{q_j m_j m_{j-1} \cdots m_1}, \quad (18)$$

with $I_{m'}$ being the identity operator on the m' components of the environment, see Fig. 1 b).

An alternative, but fully equivalent, representation for non-anticipatory channels is obtained by adding to Eq. (17) a collection of local environments which individually couples with the carriers, i.e.

$$\Phi^{(n)}(\rho_Q^{(n)}) = \text{Tr}_{ME} \left[U_{q_n M e_n} \cdots U_{q_1 M e_1} (\rho_Q^{(n)} \otimes \Omega_M \otimes \omega_E^{\otimes n}) U_{q_1 M e_1}^\dagger \cdots U_{q_n M e_n}^\dagger \right], \quad (19)$$

where for all $j = 1, 2, \dots, n$, $U_{q_j M e_j}$ is now the unitary transformation which describes the coupling of the j -th carrier with its own local environment e_j and with the memory system M , where $\omega_E^{\otimes n} := \omega_{e_1} \otimes \cdots \otimes \omega_{e_n}$ as in the memoryless case, and Ω_M is some fixed state of M , see Fig. 1 c). Equation (19) was first introduced in Ref. Bowen and Mancini, 2004 as a model for representing correlated channels: from Eq. (17) it follows that it provides a general unitary dilation for *every* non-anticipatory quantum maps. It can also be expressed in terms of a n -fold concatenation of a sequence of CPTP maps acting on a single carrier *and* the memory system M (Bowen and Mancini, 2004; Kretschmann and Werner, 2005). Such concatenation is shown pictorially in Fig. 1 c) and results in the following identity

$$\Phi^{(n)}(\rho_Q^{(n)}) = \text{Tr}_M \left[\Phi_{QM}^{(n)}(\rho_Q^{(n)} \otimes \Omega_M) \right], \quad (20)$$

with

$$\Phi_{QM}^{(n)} := \Phi_{q_n M} \circ \Phi_{QM}^{(n-1)} = \Phi_{q_n M} \circ \cdots \circ \Phi_{q_1 M}, \quad (21)$$

where for $j = 1, \dots, n$, $\Phi_{q_j M} : \mathfrak{S}(\mathcal{H}_{q_j} \otimes \mathcal{H}_M) \rightarrow \mathfrak{S}(\mathcal{H}_{q_j} \otimes \mathcal{H}_M)$ is a CPTP map that operates on the j -th carrier and on the memory ancilla M and which is defined by the unitary dilation

$$\Phi_{q_j M}(\cdots) = \text{Tr}_{e_j} \left[U_{q_j M e_j} (\cdots \otimes \omega_{e_j}) U_{q_j M e_j}^\dagger \right]. \quad (22)$$

Cases of special interest (Kretschmann and Werner, 2005) are those in which, for all j , the $\Phi_{q_j M}$ describes the same mapping $\Phi = \Phi_{qM}$ on $\mathfrak{S}(\mathcal{H}_q \otimes \mathcal{H}_M)$ which, according to Eq. (21) becomes the *generator* of the n -fold concatenation. Memoryless channels can then be included in this class as limiting case in which the generator Φ can be expressed as a tensor product channel that acts independently on the carrier q and on the memory system M . In terms of the unitary dilation (19) this is equivalent to assume that the unitaries $U_{q_j M e_j}$ in Eq. (14) factorize in a tensor product $U \otimes V_M$, where V_M is a unitary operator on the memory system and U is a unitary acting only on the degree of freedom of the j -th carrier and on its local environment e_j .

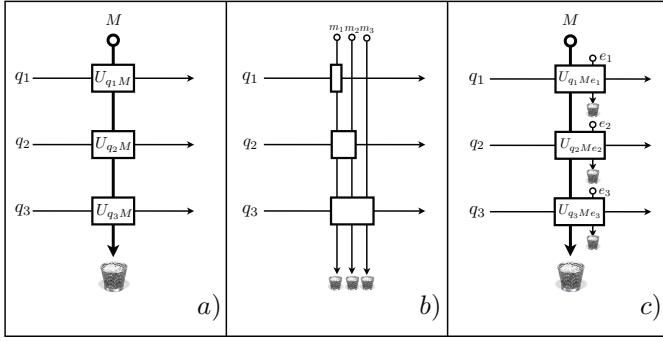


FIG. 1 Unitary dilations for a non-anticipatory quantum memory channels. a) graphical sketch of the representations of Eq. (17): here the noise correlations among the n channel uses can be described via a series of concatenated unitary interactions with a common reservoir M whose dimension in general depends (exponentially) upon n ($n = 3$ in the example). Notice that while the carrier q_1 might influence the outcome of q_2q_3 via their common interaction with M , q_2q_3 cannot influence the output of the first carrier; b) the environment M can be also represented as a collection of smaller systems M_1, M_2, \dots initially prepared into a separable state while, as shown in Eq. (18), the unitary transformation operating on the j -th channel use couples it with the first j subsystems only; c) unitary dilation (19) where apart from M a series of local environment e_1, e_2, \dots , are also present. In all the diagrams the unitary operators (represented by the white boxes) are applied sequentially on the input states of the global system (i.e., the carriers and the environment) starting for the one on the top of the figure. The carriers and the environmental states evolve, respectively, from-left-to-right and from-top-to-bottom while interacting meeting at a white box. The trash-bin symbol stands for the partial trace operation on the environment.

2. Finite-memory channels

The word *finite-memory* (Bowen and Mancini, 2004) has been used to indicate those non-anticipatory channels which admit a representation of the form (19) with M being finite dimensional. The dimension of the memory is determined by the number of Kraus operators in the single channel expansion and the correlation length of the channel, which may be defined as the maximum number of channel uses for which the noise is not conditionally independent. Any channel with a finite correlation length may be generated by a channel with a finite memory, according to this model. Within the representation (17) examples of finite-memory channels are obtained by assuming that the unitary transformations (18) couple the carriers with no more than a fixed number k of environmental subsystems, the parameter k playing the role of the correlation length of the channel. More precisely for all $j \geq k$ one has,

$$U_{q_j M} = I_{m_n} \otimes \dots \otimes I_{m_{j+1}} \otimes U_{q_j m_j m_{j-1} \dots m_{j-k}} \otimes I_{m_{j-k-1}} \otimes \dots \otimes I_{m_1}, \quad (23)$$

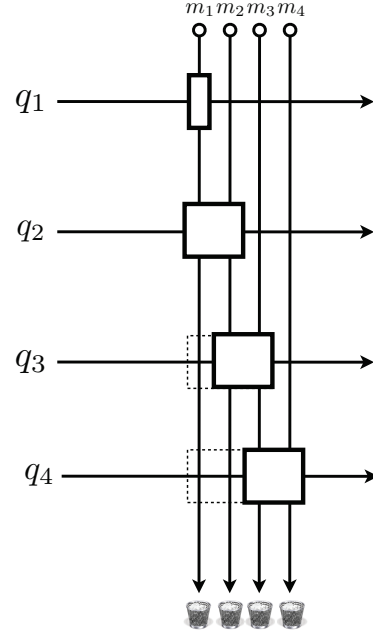


FIG. 2 Unitary dilation (17) for a finite-memory non-anticipatory quantum memory channel with correlations length $k = 2$. In the depicted example the total channel uses is $n = 4$ and each carrier is supposed to interact with only two components of the environment. Consequently the carrier q_1 can influence the output carrier q_3 only via q_2 . For a comparison see the scheme of Fig. 1 b) where instead the first carrier can directly influence q_3 via their common interaction with m_1 .

with (see Fig. 2 for a graphical representation of the case with $k = 2$). Extreme case of finite-memory channels is provided by the memoryless one where $k = 1$ and each carrier interacts with a devoted component of the multipartite environment M (specifically, for each j , the carrier q_j interacts with m_j only).

3. Symbol independent vs intersymbol interference maps

A special subset of the non-anticipatory channels is formed by the *symbol independent* (SI) maps. They are communication lines where previous input states *do not* affect the action of the channel on the current input state. In other words the symbol independent maps are non-anticipatory (or semicausal) with respect to *all* possible ordering of the carriers. Therefore given a generic subsets of the carrier set Q , its output state is uniquely determined by the corresponding input state via a proper CPTP mapping. Following the notation introduced in Refs. (Beckman *et al.*, 2001; Eggeling, Schlingemann and Werner, 2001; Piani *et al.*, 2006), they can be said *non-signaling* (or *causal*) channels, meaning that the output states of any subset of the carriers cannot be influenced by the input state of the remaining carriers.

This is opposed to the channels with *intersymbol interference* (ISI), where the input states of previous carriers affect the action of the channel on the current input. An extreme example of such channels is the quantum shift channel, where each input state is replaced by the previous input state.

C. Quasi-local algebras approach

Till now one has followed a constructive approach to quantum channels with memory, where CPTP maps which process long messages were always thought of as concatenations of smaller units which, starting from an official “first carrier” element, process one quantum signal each. An alternative view is proposed in Kretschmann and Werner, 2005 by treating the channels as mappings applied on an infinitely long message strings. This approach requires some advanced mathematical tools, like quasi-local algebras, that are sketched in Appendix D.

1. The Structure of Causal Channels

To set the stage, suppose to have a quantum channel which transforms input states of an infinitely extended quantum lattice system [representing the infinite message string] into output states on the same system. In (Kretschmann and Werner, 2005) this map is formally assigned by working in the Heisenberg picture via the introduction of a completely positive and unital map $\Phi^*: \mathcal{B}^{\mathbb{Z}} \rightarrow \mathcal{B}^{\mathbb{Z}}$ operating on the quasi-local algebra $\mathcal{B}^{\mathbb{Z}}$ that defines the observable quantities on the lattice. In this context one says that the channel is translational invariant if Φ^* commutes with the shift operator on the lattice. Requiring then that future inputs should not affect past measurements [non-anticipatory constraint (15)], Ref. (Kretschmann and Werner, 2005) introduces the definition of *causal channel* as a completely positive and unital translational invariant map Φ^* that fulfills the constraint

$$\Phi^* \left(b^{(-\infty, z]} \otimes \mathbb{1}^{[1+z, \infty)} \right) = \Phi^* \left(b^{(-\infty, z]} \right) \otimes \mathbb{1}^{[1+z, \infty)} , \quad (24)$$

for all $z \in \mathbb{Z}$ and $b^{(-\infty, z]} \in \mathcal{B}^{(-\infty, z]}$. Example of causal maps are provided by *concatenated memory channels* (Kretschmann and Werner, 2005) which can be easily constructed by adapting the concatenation scheme of Eqs. (20)-(21) to the quantum lattice formalism. Interestingly enough Ref. (Kretschmann and Werner, 2005) proves a structure theorem similar to the one discussed in Sec. III.B.1, which shows that *any* causal map (24) can always be represented as concatenated memory channel with a uniform generator.

2. Ergodic channels with decaying input memory

The notion of channels with memory having a nonanticipatory and decaying character can be extended to quantum channels having ergodic properties. Actually such properties have been established till now (Bjelaković and Boche, 2008) only for classical-quantum channels, i.e. CPTP maps that converts classical inputs into states of a quantum system (Holevo and Giovannetti, 2012).

To give an idea of how things work for classical-quantum channels (cq-channels), let A be a finite set and let \mathcal{H} be a d -dimensional Hilbert space. By $A^{\mathbb{Z}}$ one denotes the set of doubly infinite sequences with components from A and $\mathcal{B}^{\mathbb{Z}}$ is the quasi-local C^* -algebra corresponding to the set of bounded operators \mathcal{B} on \mathcal{H} (see Appendix D). Then, a cq-channel is usually meant as a triple (A, \mathcal{H}, W) , where W maps elements of A into density operators on \mathcal{H} . By the quasi-local algebra formalism it can be regarded as the map $W : A^{\mathbb{Z}} \times \mathcal{B}^{\mathbb{Z}} \rightarrow \mathbb{C}$ such that for each $b \in \mathcal{B}^{\mathbb{Z}}$, $W(x, b)$ is measurable and for each $x \in A^{\mathbb{Z}}$, $W(x, b)$ is a state.

A cq-channel W is said to be *stationary* if $W(T_{\text{in}}x, b) = W(x, T_{\text{out}}b)$ holds true for all $x \in A^{\mathbb{Z}}$ and all $b \in \mathcal{B}^{\mathbb{Z}}$. Here T_{in} and T_{out} denote the shift on the input and output algebra respectively. A cq-channel W is called *ergodic*, if it is extremal in the convex set of stationary cq-channels. As such it cannot be written as a statistical mixture of cq-channels $W = \sum_i p_i W_i$. Following the definition of causality given in Section III.C.1, a cq-channel W is causal if for each $n \in \mathbb{Z}$, $b \in \mathcal{B}^{(-\infty, n]}$ and all $x, \tilde{x} \in A^{\mathbb{Z}}$ ($x_i = \tilde{x}_i; \forall i \leq n$), it is

$$W(x, b) = W(\tilde{x}, b). \quad (25)$$

Furthermore, W is called *input memoryless* if for each $n \in \mathbb{Z}$, $b \in \mathcal{B}^{[n, \infty)}$ and all $x, \tilde{x} \in A^{\mathbb{Z}}$ ($x_i = \tilde{x}_i; \forall i \geq n$) it is

$$W(x, b) = W(\tilde{x}, b). \quad (26)$$

In the presence of memory it is expected that the effect of the far past inputs do not affect much present and future outputs. According to this idea, a channel W is said to have *decaying input memory (and anticipation)* if for each $\epsilon > 0$ there exist non-negative integers $m(\epsilon), a(\epsilon)$ such that

$$|W(x, b) - W(x', b)| \leq \epsilon, \quad (27)$$

for all $b \in \mathcal{B}^{[n, n+k]}$, $n, k \in \mathbb{Z}$, whenever $x_i = x'_i$ for $n - m \leq i \leq n + k + a$ and $m \geq m(\epsilon)$, $a \geq a(\epsilon)$. This provides a ‘continuity’ property of the channel which plays a crucial role in establishing coding theorems, an idea that will also appear in Section IV.D and goes back to the classic paper by McMillan (McMillan, 1953).

IV. TAXONOMY OF QUANTUM MEMORY CHANNELS

This section focuses on some classes of non-anticipatory memory quantum channels which have attracted special interest in the literature.

A. Localizable, fully non-anticipatory memory quantum channels

A subset of non-anticipatory quantum channels which have been extensively analyzed (Giovannetti and Mancini, 2005; Plenio and Virmani, 2007, 2008) is graphically described in Fig. 3. For such models, the mapping (15) can be expressed in terms of (not necessarily identical) *local* unitary couplings with a correlated many-body environmental system $E := \{e_1, e_2, \dots\}$. Memory effects in the communication appear because, differently from the memoryless case (14), the many-body environment is not initialized in a state $\omega_E^{(n)}$ which does not factorize, i.e.,

$$\Phi^{(n)}(\rho_Q^{(n)}) = \text{Tr}_E \left[U_{q_n e_n} \otimes \dots \otimes U_{q_1 e_1} \left(\rho_Q^{(n)} \otimes \omega_E^{(n)} \right) U_{q_1 e_1}^\dagger \otimes \dots \otimes U_{q_n e_n}^\dagger \right]. \quad (28)$$

An alternative representation of these maps has been provided in Ref. (Caruso, Giovannetti and Palma, 2010) by generalizing a model presented in Ref. (Ban, Sasaki and Takeoka, 2002; Bowen and Bose, 2004) for memoryless channels. According to this approach the channel noise is effectively described as a quantum teleportation protocol (Bennett *et al.*, 1993; Braunstein and Kimble, 1998; Vaidman, 1994) that went wrong because the communicating parties used non optimal resources (e.g., the state they shared was not maximally entangled). In the case of (28) each of the carriers gets teleported independently using the same procedure, the correlations arising from the fact that the communicating parties use as shared resource a correlated many-body quantum state.

One can easily verify that the maps described by Eq. (28) are symbol independent. More precisely in the language of Refs. (Beckman *et al.*, 2001; Egeling, Schlingemann and Werner, 2001; Piani *et al.*, 2006) they are *localizable*, *non-signaling* channels, since by construction admit a representation in terms of local coupling with a correlated ancilla (it is worth reminding that *not all* the non-signaling maps can be represented in this form, see e.g. (Beckman *et al.*, 2001; Piani *et al.*, 2006; Popescu and Rohrlich, 2001)).

B. Perfect memory channels

The memoryless channels are characterized by possessing unitary dilations in which the environment has a dimension which is exponentially growing in n (i.e., $\log[\dim \mathcal{H}_E^{(n)}] = n \log[d_e]$) or, equivalently, by possessing a (minimal) operator sum representations whose Kraus sets contains a number of elements which is exponentially growing in n . This same property typically holds also for memory channels with the important exception of the *perfect memory* channels (Giovannetti, Burgarth and Mancini, 2009; Kretschmann and Werner, 2005). The simplest example of such communication lines is obtained by assuming that the memory system M in Eq. (17) does

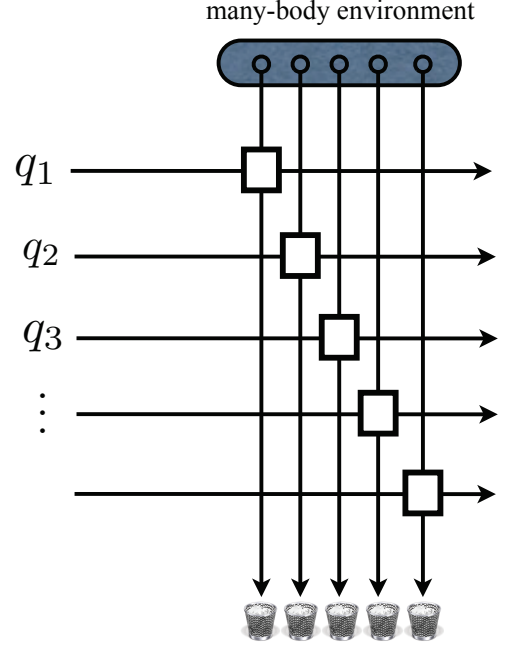


FIG. 3 Model for a localizable, fully non-anticipatory quantum memory channel. Here the correlations are introduced by allowing the state of the environment (gray element) to be initially entangled. As in the previous figures white boxes represents unitary couplings while the trash-bin indicates partial trace over the corresponding degree of freedom. These maps are SI and hence non-anticipatory (therefore they also admit unitary dilations of the form described in Fig. 1).

not scale with n and it is finite dimensional. Under this hypothesis the maps $\Phi^{(n)}$ explicitly admit a unitary dilation with an environment (the system M) of constant size which is hence unable to capture all the information that is sent through the channel (the latter being measured by the size of the carriers $\mathcal{H}_Q^{(n)}$ which is exponential in n). As a consequence, in the asymptotic limit of long carrier sequences, no information is expected to be lost to the environment, yielding optimal quantum communication efficiency. A typical example is provided by the shift channel (see Sec. III.B.3) which can be described as in Eq. (17) by assuming M to have the same dimension of a single carrier and by taking U_{qjM} as the SWAP operator. It is worth noticing that in Ref. (Bowen, Devetak and Mancini, 2005) it was also conjectured that the types of memory channels that, analogously to the shift channel, display *only* intersymbol interference may be represented as perfect memory channels.

More generally the class of perfect memory channels can be extended to include all the CPTP maps (15) that admit unitary dilations (17) in which the dimension d_M of the environmental system M is sub-exponential in n , i.e.,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log[d_M] = 0. \quad (29)$$

As discussed explicitly in Sec. VI.B.2, also in this case the channel capacity of the associated communication line is optimal (Giovannetti, Burgarth and Mancini, 2009; Kretschmann and Werner, 2005).

C. Markovian channels

An important class of non-anticipatory quantum channels is given by the channels with Markovian correlated noise. They describe noise models in which the carriers are transformed via the applications of strings of local CPTP maps whose elements are randomly generated by a classical Markov process. For instance a quantum Markov process is given by

$$\Phi^{(n)}(\rho_Q^{(n)}) = \sum_{i_1, \dots, i_n} p_{i_n|i_{n-1}}^{(n)} p_{i_{n-1}|i_{n-2}}^{(n-1)} \cdots p_{i_2|i_1}^{(2)} p_{i_1}^{(1)} \times \Phi_{q_n}^{(i_n)} \circ \Phi_{q_{n-1}}^{(i_{n-1})} \circ \cdots \circ \Phi_{q_1}^{(i_1)}(\rho_Q^{(n)}), \quad (30)$$

where $\{\Phi_{q_j}^{(i)}\}_i$ is a set of CPTP maps operating on the j -th carrier, where $p_i^{(1)}$ is an initial probability distribution, and where finally for $j \geq 2$, the $p_{i|i'}^{(j)}$ are conditional probabilities.

The mapping (30) is symbol independent since modifying the input state of previous channels uses do not have any effects on the output states of the carriers that follows. A unitary dilation of the form (19) can be obtained by identifying the initial state Ω_M of the memory M with the pure vector $\sum_i \sqrt{p_i^{(1)}}|i\rangle_M$, and by taking the unitary $U_{q_j M e_j}$ in such a way that for all vectors $|\psi\rangle_{q_j}$ of the j -th carriers one has

$$U_{q_1 M e_1}|\psi, i', 0\rangle = A_{q_1}^{(i')}(\ell)|\psi, i', \phi_{\ell}'\rangle, \quad (31)$$

for $j = 1$, while

$$U_{q_j M e_j}|\psi, i', 0\rangle = \sum_i \sqrt{p_{i|i'}^{(j)}} A_{q_j}^{(i)}(\ell)|\psi, i, \phi_{\ell}'\rangle, \quad (32)$$

for $j \geq 2$ (in the above expressions $\{A_{q_j}^{(i)}(\ell)\}_{\ell}$ is a Kraus set for $\Phi_{q_j}^{(i)}$, $|\psi, i, \phi_{i', \ell}\rangle$ stands for the state $|\psi\rangle_{q_j}|i\rangle_M|\phi_{i', \ell}\rangle_{e_j}$, while $\{|i\rangle_M\}_k$ and $\{|\phi_{i, \ell}\rangle_{e_j}\}_{i, \ell}$ are orthonormal basis for the memory systems M and e_j respectively).

Most of the analysis conducted so far focused on the special case of homogeneous Markov processes in which both the $p_{i|i'}^{(j)}$ and the $\Phi_{q_j}^{(i)}$ do not depend upon the carrier label j (e.g., $p_{i|i'}^{(j)} := p_{i|i'}$). Under these conditions one also says that the quantum Markov process is *regular* if the corresponding classical Markov process $p_{i|i'}$ is regular, i.e., if some power of the transition matrix Γ of elements $p_{i|i'}$ has only strictly positive elements. In this case, for $j \rightarrow \infty$ the statistical distribution of the local noise converges to a stationary distribution

$p_i^{(\infty)} := \lim_{n \rightarrow \infty} p_i^{(n)}$ with

$$p_i^{(n)} := \sum_{i'} (\Gamma^{n-1})_{i, i'} p_{i'}^{(1)}, \quad (33)$$

being the probability of getting $\Phi_{q_j}^{(i)}$ on the j -th carrier. Also the initial probability $p_i^{(1)}$ is said to be stationary if it satisfies the eigenvector equation $\sum_{i'} \Gamma_{i, i'} p_{i'}^{(1)} = p_i^{(1)}$ (when this happens $p_i^{(j)} = p_i^{(1)}$ and the local statistical distribution of $\Phi_{q_j}^{(i)}$ is identical for all the carriers).

The first example of a regular Markov process has been analyzed by Macchiavello and Palma, 2002. Here the carriers are assumed to be qubits and the CPTP transformations $\Phi_q^{(i)}$ entering in Eq. (30) perform unitary rotations $\Phi_q^{(i)}(\cdots) := \sigma_q^{(i)}(\cdots)\sigma_q^{(i)}$ where $\sigma_q^{(0)} = I$ is the identity operator while for $i = 1, 2, 3$, $\sigma_q^{(i)}$ is the i -th Pauli spin matrix. The conditional probability $p_{i|i'}$ which describes the associated classical Markov process was finally written as $p_{i|i'} = (1 - \mu)p_i^{(1)} + \mu\delta_{ii'}$ where $\mu \in [0, 1]$ is a correlation parameter (notice that for $\mu = 0$ the model describes a memoryless channel while for $\mu = 1$ it describes a long-memory channel – see Sec IV.E). This model of Markovian correlated Pauli channel show a remarkable feature when it is used for the transmission of classical information (see Sec. VI). That is, when two successive uses of the channel are considered, classical information is optimally encoded in either separable states or maximally entangled states, depending whether the correlation parameter μ is below or above a certain threshold value. This feature was first conjectured in (Macchiavello and Palma, 2002), then proven for certain instances of the model in (Macchiavello, Palma and Virmani, 2004), and finally proven for general Markovian correlated Pauli channel in (Daems, 2007).

An experimental demonstration of the optimality of entangled qubit pairs for encoding classical information through a correlated Pauli channel was provided by Banaszek *et al.*, 2004 for mechanically induced correlated birefringence fluctuations, which in turn induce correlated depolarization (Ball, Dragan and Banaszek, 2004).

A generalized model of d -dimensional Markovian correlated Pauli channel was considered by Shadman *et al.*, 2011 for the problem of sending classical information using a dense-coding protocol. An alternative model of two-qubit correlated channel was characterized by Caruso *et al.*, 2008b in terms of the minimum output entropy.

Going beyond the case of two uses of a qubit channel, Markovian correlated depolarization over an arbitrary number of channel uses was studied in (Demkowicz-Dobrzanski, Kolenderski and Banaszek, 2007; Karimipour and Memarzadeh, 2006b), and the case of Markovian correlated noise in higher dimensional quantum systems was considered in (Karimipour and Memarzadeh, 2006; Karpov, Daems and Cerf, 2006,b). Generally speaking, the optimality of entangled state encoding can be interpreted as a consequence of a decoherence-free subspace (see Sec. V) associated to the correlated noise

model: this has been considered for the Hilbert space defined by multiple uses of a qubit channel in (Demkowicz-Dobrzanski, Kolenderski and Banaszek, 2007) and for the multiphoton Hilbert space associated to the polarization of light (Ball and Banaszek, 2005). In a different context, the same phenomenon has been discussed for the problem of quantum communication with polarized light without a shared reference frame (Bartlett, Rudolph and Spekkens, 2003).

Finally, models of Markovian correlated noise in the framework of quantum systems with continuous variable (see Sec. VII.B) were first discussed in (Cerf *et al.*, 2005, 2006) for the case of two uses of the channel, and then extended to the arbitrary number of uses in (Lupo, Memarzadeh and Mancini, 2009; Schäfer, Karpov and Cerf, 2009) (see Sec. VII.B).

D. Fixed-point, indecomposable, and forgetful channels

Memory channel representations where the action on the memory state is independent of the input state are termed *fixed-point* memory channels, as the memory state will have a fixed point under the action of the representation (Bowen, Devetak and Mancini, 2005). Specifically remembering the definitions of Eqs. (20) and (21) this implies,

$$\Omega_M(\rho, n) := \text{Tr}_Q \left[\Phi_{QM}^{(n)}(\rho_Q^{(n)} \otimes \Omega_M) \right] = \Omega_M, \quad (34)$$

for all n and for all $\rho_Q^{(n)}$. Fixed-point channels can be easily shown to be symbol independent while the opposite is not necessarily true. Indeed from Eq. (21) one has that the output state of n -th carrier $\rho'_{q_n} := \text{Tr}_{Q^{(n-1)}} \left[\Phi^{(n)}(\rho_Q^{(n)}) \right]$ can be expressed as

$$\begin{aligned} \rho'_{q_n} &= \text{Tr}_{Q^{(n-1)}M} \left[\Phi_{q_n M} \circ \Phi_{QM}^{(n-1)}(\rho_Q^{(n)} \otimes \Omega_M) \right] \\ &= \text{Tr}_M \left[\Phi_{q_n M}(\rho_{q_n} \otimes \Omega_M) \right], \end{aligned} \quad (35)$$

which only depends upon the reduced density matrix ρ_{q_n} (in these expression $\text{Tr}_{Q^{(n-1)}M}[\cdots]$ indicates the partial trace with respect to M and the first $(n-1)$ carriers).

An *indecomposable* channel is one where, for each channel input, the long-term behavior of the channel is independent of the initial memory state (Bowen, Devetak and Mancini, 2005). More precisely a finite-memory quantum channel is indecomposable if for any input state ρ and $\epsilon > 0$ there exists an $N(\epsilon)$ such that for $n \geq N(\epsilon)$,

$$D(\Omega_M(\rho, n), \Sigma_M(\rho, n)) \leq \epsilon, \quad (36)$$

where $\Omega_M(\rho, n)$ and $\Sigma_M(\rho, n)$ are the memory states after n uses of the channel for the initial memory states Ω_M and Σ_M respectively, and where $D(\rho, \rho')$ is the trace distance (see Appendix A).

The main features of indecomposable channels have been revisited through the notion of *forgetful channels* (Kretschmann and Werner, 2005) which it is presented here in a slightly more general form. Let Φ_{qM} be a CPTP map on $\mathfrak{S}(\mathcal{H}_q \otimes \mathcal{H}_M)$ which generates a non-anticipatory channel $\Phi_{QM}^{(n)}$ via the n -fold concatenation, Eqs. (20)-(22). We say that Φ_{qM} is forgetful iff there exists a sequence of quantum channels $S_n : \mathfrak{S}(\mathcal{H}_q \otimes \mathcal{H}_M) \rightarrow \mathfrak{S}(\mathcal{H}_M)$ such that

$$\lim_{n \rightarrow \infty} \|\Phi_{QM}^{*(n)} - \text{id}_M \otimes S_n^*\|_{cb} = 0, \quad (37)$$

where $\|\cdots\|_{cb}$ is the cb-norm (see Appendix A) while $\Phi_{QM}^{*(n)}$ and S_n^* are the Heisenberg dual of the maps $\Phi_{QM}^{(n)}$ and S_n respectively. An example is given by the channel $\Phi_{QM}^{(n)}$ obtained via concatenation (22) using as generator the map $\Phi_{qM} := p \text{id} + (1-p) \text{SWAP}$, where $p \in [0, 1]$, and SWAP denotes the swap channel which exchanges q and M . The only way for $\Phi_{QM}^{(n)}$ not to be forgetful is to choose the ideal channel in every step. However, the probability for this event vanishes in the limit $n \rightarrow \infty$ as p^n , implying that Eq. (37) holds.

The equivalence between the notion of indecomposable channels and that of forgetful channels has been shown in (Kretschmann and Werner, 2005). Several criteria for a quantum memory channel to be forgetful exist [for instance, a sufficient condition is to have that $\|\Phi_{QM}^{*(n)} - \text{id}_M \otimes S_n^*\|_{cb}$ falls below 1 for *some* finite n].

From a physical point of view, one could expect a generic quantum memory channel to be forgetful. Indeed, it can be proven that the subset of forgetful channels is dense and open (according to the topology induced by the norm of complete boundness) (Kretschmann and Werner, 2005). In the case of Markovian channels, the forgetfulness is determined by the asymptotic properties of the underlying Markov chain: in particular, for a discrete variable memory system, the channel is forgetful if and only if the underlying Markov chain converges to a unique stationary state (Datta and Dorlas, 2009). On the other hand, if the memory system is described by continuous variables, one could have situations in which the Markov chain has a unique stationary state, yet the convergence property (37) is not verified. To overcome this limitation, a weaker notion of forgetfulness, named *weak forgetfulness*, has been introduced in (Lupo, Memarzadeh and Mancini, 2009). This notion applies only to fixed-point memory channels with memory systems described by classical degrees of freedom. Although restricted to this setting, its definition coincides with that of indecomposability [Eq. (36)], and is equivalent to forgetfulness for discrete variable Markov chain. It has also been argued in (37), but not proven rigorously, that the coding theorem for forgetful channels (see Sec. VI.B.4) can be extended to weak forgetful ones. Beyond this setting, the model of Gaussian memory channel in (Lupo, Giovannetti and Mancini, 2010) was proven to be forgetful under conditions on the memory initialization, e.g., if the

initial state of the memory is a Gaussian state with finite first and second moments. Finally, the relation between the forgetfulness of the channel and the chaotic quantum evolution of its environment was studied in (Barreto Lemos and Benenti, 2010) for the case of a model of dephasing channel with memory.

E. Long term memory channels

Long term quantum memory channels describe those communication lines in which the effect of the memory does not die with the channel uses. Extreme examples are provided by statistical mixture of memoryless channels (Bjelaković and Boche, 2009; Bjelaković, Boche and Nötzel, 2009; Datta and Dorlas, 2007; Datta, Suhov and Dorlas, 2008) where the carriers are transformed according to the mapping

$$\Phi^{(n)}(\rho_Q^{(n)}) = \sum_i p_i \Phi_i^{\otimes n}(\rho_Q^{(n)}), \quad (38)$$

where $\{\Phi_i\}_i$ is a set of single carrier CPTP channels, and p_i is a probability distribution on such set. Equation (38) describes a scenario in which, with some probability p_i all the carriers of the system are operated by the same identical local transformation Φ_i . The index i can be interpreted as a “switch” selecting different memoryless channels, and the memory channel in (38) can be interpreted as the average channel over different values of the “switch”.

It is worth stressing that the transformation (38) is fully non-anticipatory (i.e., the input state of any subset of carriers cannot influence the output state of the remaining ones): as a consequence, fixing an ordering, it can always be represented as in Fig. 1 with a proper choice of the unitary couplings. Also the memory correlations of this class of channels can be considered to be given by a Markov chain which is aperiodic but not irreducible (Norris, 1997). This can be easily seen by noticing that Eq. (30) reduces to Eq. (38) by assuming the $\Phi_{qj}^{(j)}$ appearing in the former expression not to depend upon the carrier and by setting $p_{i|i'}^{(j)} = \delta_{ii'}$ for all $j = 2, \dots, n$. Hence, once a particular branch, $i = 1, \dots, M$, has been chosen, the successive inputs are sent through this branch: transition between the different branches (which correspond to the different states of the Markov chain) is not permitted.

V. QUANTUM CODES

The channel formalism introduced in the previous sections provides the tools for describing the effects of noise on a quantum system. Within this formalism the Kraus operators have to be intended as errors introduced onto the message represented by the system’s (input) state. Given that one aims at transmitting (classical or quantum) information through the quantum channel, the

presence of noise may impose severe limitations to the *reliable* information transmission. The key idea, mutated from classical information theory, to realize reliable information transmission in the presence of noise is to exploit redundancy by means of a suitable *quantum error correcting code* (QECC). The latter must be designed to make the received noisy signal decodable with an asymptotically negligible probability of error.

While classical works on QECC were mainly concerned in counteracting noise arising from memoryless quantum channels, the effects of noise correlations and memory characterizing quantum memory channels posed new challenges for the optimal design of QECC.

More specifically, one says that information can be reliably sent through a quantum channel $\Phi^{(n)} : \mathfrak{S}(\mathbb{C}^{2^{\otimes n}}) \rightarrow \mathfrak{S}(\mathbb{C}^{2^{\otimes n}})$ if there exist an encoding map

$$\Phi_E^{(k \rightarrow n)} : \mathfrak{S}(\mathbb{C}^{2^{\otimes k}}) \rightarrow \mathfrak{S}(\mathbb{C}^{2^{\otimes n}}), \quad (39)$$

with $n \geq k$, and a decoding map

$$\Phi_D^{(n \rightarrow k)} : \mathfrak{S}(\mathbb{C}^{2^{\otimes n}}) \rightarrow \mathfrak{S}(\mathbb{C}^{2^{\otimes k}}), \quad (40)$$

such that $\forall |\psi\rangle \in \mathfrak{M}$ it is

$$\lim_{k \rightarrow \infty} F(|\psi\rangle\langle\psi|, \Phi_D^{(n \rightarrow k)} \circ \Phi^{(n)} \circ \Phi_E^{(k \rightarrow n)}(|\psi\rangle\langle\psi|)) = 1. \quad (41)$$

Here \mathfrak{M} is the message space corresponding to $\{|0\rangle, |1\rangle\}^{\otimes k}$ in case of classical information transmission, and to $\mathbb{C}^{2^{\otimes k}}$ in case of quantum information transmission. Furthermore, in Eq. (41) F is the fidelity (Jozsa, 1994; Uhlmann, 1976), defined as in Eq. (A3). With the above conditions fulfilled, one says that $R = k/n$ is a *reliable* communication rate. We will address in the next Section the question of what are the maximum reliable rates for a given quantum channel. Here, one is interested on the image of the encoding map $\Phi_E^{(k \rightarrow n)}$, denoted $\mathcal{C} \subset \mathbb{C}^{2^{\otimes n}}$, which is called *quantum error correcting code* (QECC). The state of k qubits is encoded in \mathcal{C} and has to be decoded from \mathcal{C} with asymptotically unit fidelity.

A. Standard quantum coding theory

Let us consider a quantum memoryless channel $\Phi^{(n)}$, characterized by Kraus operators $\{A_i\}$, on the Hilbert space $\mathbb{C}^{2^{\otimes n}}$ of n qubits

$$\rho \mapsto \Phi^{(n)}(\rho) = \sum_i A_i \rho A_i^\dagger, \quad (42)$$

where $A_i = A_{i_1} \otimes \dots \otimes A_{i_n}$ describes independent and identically distributed errors on single qubits. A $[[n, k, d]]$ quantum correcting code \mathcal{C} is given by a 2^k -dimensional subspace of $\mathbb{C}^{2^{\otimes n}}$ where it is possible to correct all errors affecting at most $t = \lfloor (d-1)/2 \rfloor$ qubits, i.e., corresponding to those Kraus operators having at most t operators A_{i_j} ($j = 1, \dots, n$) different from identity. Furthermore, one denotes by $P_{\mathcal{C}}$ the projector onto the quantum code

\mathcal{C} , and by \mathfrak{E} a subset of the error operators $\{A_i\}$. The quantum code \mathcal{C} is able to correct all errors in \mathfrak{E} if and only if there exists a hermitian matrix γ such that

$$P_{\mathcal{C}} A_i^\dagger A_{\mathbf{m}} P_{\mathcal{C}} = \gamma_{\mathbf{lm}} P_{\mathcal{C}}, \quad (43)$$

for any pair of error operators $A_{\mathbf{l}}, A_{\mathbf{m}} \in \mathfrak{E}$ (Knill, 2002). The pair $(\mathcal{C}, \mathfrak{E})$, consisting of a quantum code \mathcal{C} and a vector space of error operators \mathfrak{E} , is called *degenerate* if the matrix γ in (43) is singular; otherwise, $(\mathcal{C}, \mathfrak{E})$ is said to be *nondegenerate*. By exploiting the symmetry of the map under isometry transformation of the Kraus operators (see Sec. II), it follows that in the nondegenerate case different error operators map the code in orthogonal subspaces, while for degenerate codes it may happen that distinct error operators transform the code into non-orthogonal subspaces.

The action of the correctable error operators \mathfrak{E} on vectors spanning \mathcal{C} (codewords) allows to define the CPTP recovery map $\Phi_R^{(n)} \leftrightarrow \{R_{\mathbf{l}}\}$. The composition of the recovery map $\Phi_R^{(n)}$ with the quantum channel gives the recovered channel,

$$\Psi(\rho) := \left(\Phi_R^{(n)} \circ \Phi^{(n)} \right)(\rho) = \sum_{\mathbf{l}} \sum_{\mathbf{k}} (R_{\mathbf{l}} A_{\mathbf{k}}) \rho (R_{\mathbf{l}} A_{\mathbf{k}})^\dagger. \quad (44)$$

The action of the recovery map $\Phi_R^{(n)}$ may be interpreted as a unitary transformation involving an ancillary system, followed by a measurement on such ancilla aimed to extract the error syndrome and by a subsequent error correction action on the system.

An upper bound on the rates achievable by nondegenerate quantum codes is known to be the quantum version of the Hamming bound (Ekert and Macchiavello, 1996)

$$2^k \sum_{i=0}^t 3^i \binom{n}{i} \leq 2^n. \quad (45)$$

In the framework of independent and identical distributed (i.i.d.) errors, no codes are known to violate such a bound. There are also upper bounds that apply to all quantum codes, not just non-degenerate ones, like quantum Singleton bound $n \geq 4t + k$ (Knill and Laflamme, 1997).

Unfortunately the explicit construction of practical codes is not an easy task. Historically the first attempts were made following classical linear codes (Peterson and Weldon, 1972). A great advantage of linear codes over general error correcting codes is their compact specification. Then, along this line, quasi classical [or CSS (Calderbank, Shor, Steane)] codes have been proposed (Calderbank and Shor, 1996; Steane, 1996). Another way to construct quantum codes is to exploit the group structure of the set of errors \mathfrak{E} , as it is done for stabilizer codes (Gottesman, 1997). Also this kind of codes have a compact description which is mirrored into an efficient

encoding and decoding procedure. An example is provided by the $[[5, 1, 3]]$ code introduced in (Laflamme *et al.*, 2009) saturating the quantum Hamming bound (45).

The above methods of constructing codes can be summarized as follow. Consider a set \mathfrak{E} of 2^{n-k} commuting error operators of the kind $e_i = \sigma_x^{(\alpha_i)} \sigma_z^{(\beta_i)}$, ($i = 1, \dots, 2^{n-k}$), with $\alpha_i, \beta_i \in \{0, 1\}^n$ and with $\sigma_{x,y,z}$ representing the Pauli operators. The set of vectors stabilized by the error operators forms the quantum code, namely

$$\mathcal{C} = \{ |x\rangle \in \mathbb{C}^{2^n} \mid e_i |x\rangle = |x\rangle, \forall i = 1, \dots, 2^{n-k} \}. \quad (46)$$

The error operators can be represented by $2n$ dimensional vectors $v_{e_i} = (\alpha_i | \beta_i)$ in the field \mathbb{F}_2^{2n} . Let us take $n - k$ such vectors named g^1, \dots, g^{n-k} that are linearly independent and write down the following $(n - k) \times 2n$ (parity check) matrix

$$H := \begin{pmatrix} g^1 \\ \vdots \\ g^{n-k} \end{pmatrix}. \quad (47)$$

The subspace \mathcal{C} is a $[[n, k, d]]$ code, where d is the minimum Hamming distance between the vectors g^i generating the code. The quantity $H v_{e_i}^T$ represents the e_i error syndrome $\text{synd}(e_i)$. Clearly only errors e_i such that $\text{synd}(e_i) \neq 0$ can be identified and the set \mathfrak{E} is correctable by \mathcal{C} if and only if $\text{synd}(e_i) \neq \text{synd}(e_j)$ for all $i, j = 1, \dots, 2^{n-k}$ (which is the condition corresponding to (43)).

For such codes there exists a lower bound on the rates (the quantum version of the Gilbert-Varshamov bound) telling us that good codes exists (Calderbank *et al.*, 1997)

$$2^k \sum_{i=0}^{2t} 3^i \binom{n}{i} \geq 2^n. \quad (48)$$

To study the performance of a given code (even in the presence of correlated errors), it seems natural from (41) to use the fidelity (A3) as an indicator of its effectiveness. However, following Nielsen, 1996, one is led to consider entanglement fidelity (see Appendix A) as the most appropriate quantifier of the effectiveness of quantum error correcting codes (especially in the presence of correlated errors (Cafaro *et al.*, 2011)). From Eq. (A7) one obtains the entanglement fidelity associated to the error correction code:

$$F_e \left(\frac{1}{2^n} I_{\mathcal{H}}, \Psi \right) = \frac{1}{2^{2n}} \sum_{\mathbf{l}, \mathbf{k}} |\text{Tr}(R_{\mathbf{l}} A_{\mathbf{k}})|^2. \quad (49)$$

1. Performance achieved in the case of correlated errors

Although the idea of i.i.d. errors was underlying the standard theory of quantum error correcting codes, recent studies try to characterize the effect of correlated noise on the performance of most relevant QECCs.

To get some insights let us start by considering a memory channel described by the Markovian map (30) with $\Phi^{(i_l)}(\dots) = A_{i_l}(\dots)A_{i_l}^\dagger$ being the Kraus operators A_{i_l} unitary and $p_{i_l|i_{l-1}} = (1 - \mu)p_{i_l} + \mu\delta_{i_l, i_{l-1}}, \forall l = 1, \dots, n$, so to write

$$\Phi_\mu(\rho) = \sum_{i_1, \dots, i_n} p_{i_n|i_{n-1}} p_{i_{n-1}|i_{n-2}} \dots p_{i_2|i_1} p_{i_1} (A_{i_n} \otimes \dots \otimes A_{i_1}) \rho (A_{i_n} \otimes \dots \otimes A_{i_1})^\dagger. \quad (50)$$

The correlation parameter μ describes the degree of memory of the considered channel. For $\mu = 0$ one obtains the case of i.i.d. (memoryless) noise, while the limit $\mu = 1$ describes completely correlated errors.

In Ref. (Cafaro and Mancini, 2010a) the case with $A_0 = \sigma_0, A_1 = \sigma_x, \forall l, p_0 = 1 - p$ and $p_1 = p$ has been considered. Since this situation is analogous to the one with classical correlated errors, a $[n, 1, 3]$ repetition code (with $n \geq 3$) has been analyzed. It has turned out that the entanglement fidelity (49), while obviously increasing with n , is quickly decreasing as μ becomes greater than zero. The analysis was subsequently generalized (Cafaro and Mancini, 2010b) to include depolarizing quantum channels with correlated and asymmetric errors, i.e. $A_0 = \sigma_0, A_1 = \sigma_x, A_2 = \sigma_y, A_3 = \sigma_z, \forall l$ and $p_0, p_1 \neq p_2 \neq p_3$. In this case quantum error correction has been performed via the $[[5, 1, 3]]$ and $[[7, 1, 3]]$ stabilizer codes (Gottesman, 2009) showing that the performance of both codes is lowered by the presence of correlations. Furthermore, it has been uncovered that the asymmetry in the error probabilities does not affect the performance of the five-qubit code while it does affect the performance of the seven-qubit code which results less effective when considering correlated and symmetric depolarizing errors, but more effective for correlated and asymmetric errors.

Another interesting model is given by the convex combination of uncorrelated and completely correlated quantum channels (see Sec. IV.E)

$$\Phi(\rho) = (1 - \mu) \Phi_{\mu=0}(\rho) + \mu \Phi_{\mu=1}(\rho), \quad (51)$$

where $\Phi_{\mu=0}$ and $\Phi_{\mu=1}$ are given by (50) in the limiting cases of $\mu = 0$ and $\mu = 1$ (corresponding to uncorrelated and completely correlated errors respectively). For this channel, with $A_0 = \sigma_0, A_1 = \sigma_x, p_0 = 1 - p$ and $p_1 = p$, the performance of a $[n, 1, 3]$ repetition code (with $n \geq 3$) has been analyzed in Ref. (Cafaro and Mancini, 2011). Also in this case the entanglement fidelity (49), is quickly decreasing by increasing the degree of memory μ .

From the above results it seems that correlations have always a negative effects on the standard QECC. However it is worth considering other situations as well. For instance, correlated errors can arise when qubits interact with a common bath. In Ref. (Klesse and Frank, 2005) a quasi classical code has been analyzed, with variable length n (number of physical qubits) and size k (number of logical qubits) using a spin-boson model consisting of

n spins describing an n -qubit register coupled to a common bosonic bath. The possibility of exchanging bath bosons between qubits gives rise to spatial and temporal correlations in the noise. The amount of such correlations can be controlled by the inter-spin distance. Also in this case quantum error correction is substantially hampered by the kind of noise-correlations captured in the model. However, one could also consider the same code when qubits all interact with a common spin bath, rather than a boson bath. In Ref. (Shabani, 2008) such a situation has been studied using a dephasing bath interaction and it has been observed that better performance can be obtained in the presence of correlated errors, depending on the timing of the recovery.

As one has already remarked, generally in the standard QECC theory single-qubit-errors where only one qubit has undergone interaction with the environment are assumed to be the most common ones. More precisely, it is assumed that the probability of k (integer $k > 1$) errors is of order ϵ^k , which is much smaller than ϵ , given that the probability ϵ of a single error is small enough. This is the actual meaning of the error independence condition. However, it should be noted that the condition of independent errors is not equivalent to the condition of independent quantum channels, in which case each qubit interacts with its own environment that do not interact among themselves. Although the independence of the qubit-environment interactions ensures the independence condition, the converse is not guaranteed. In Ref. (Hwang, Ahn and Hwang, 2001) it is shown that even if qubits do not interact independently with environments, the generated errors satisfy the independence condition up to second order, provided that quantum bits do not directly interact with each other. Generally, this no-qubits-interaction condition is assumed except for the case where two-qubit gate operation is being performed. Thus, in such cases standard QECCs work enough well. A pictorial representation of these different kinds of system-environment interaction is depicted in Fig. 4.

Let us see why a QECC that corrects one qubit error in the standard way may fail in correcting strongly correlated errors like controlled bit flip errors. Assume to have a quantum code \mathcal{C} , with orthogonal codewords $\{|\phi_i\rangle\}_{i=1}^{|\mathcal{C}|}$, that can correct any error operator belonging to \mathfrak{E} when applied to any vector $|\phi\rangle \in \mathcal{C}$. Specifically, the code works in such a way that for any noise operator $A \in \mathfrak{E}$, a corrupted codeword $|\phi'_i\rangle = A|\phi_i\rangle$ is mapped to a product state $|\phi_i\rangle \otimes |\text{synd}(A)\rangle$, where $\text{synd}(A)$ stands for the error-syndrome of A . Since $\text{synd}(A)$ does not depend on $|\phi_i\rangle$, one can correct errors applied to any linear combination $\sum_i c_i |\phi_i\rangle$ of the basis vectors $\{|\phi_i\rangle\}_{i=1}^{|\mathcal{C}|}$. However, this is no longer true for controlled bit flip errors, where the error may depend on the specific codeword $|\phi_i\rangle$. In that case the corrupted codeword reads $\sum_i c_i A_i |\phi_i\rangle$. Then, upon decoding one gets $\sum_i c_i |\phi_i\rangle \otimes |\text{synd}(A_i)\rangle$ which, by tracing out the syndrome register, gives a

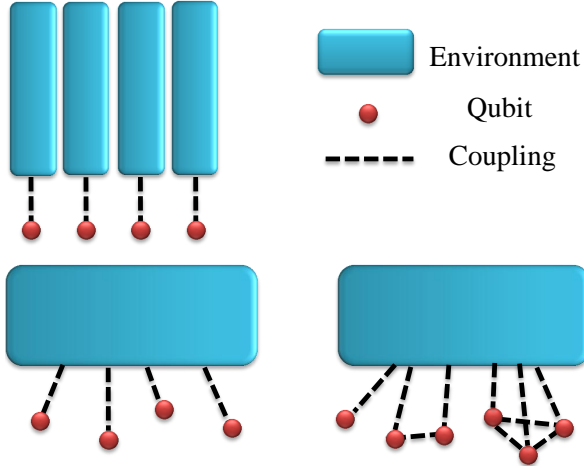


FIG. 4 A pictorial representation of different kinds of system-environment interactions.

state different of the original one $\sum_i c_i |\phi_i\rangle$. Actually, for controlled bit-flip errors, it is like the environment gets information about the codeword and thus corrupt it. This kind of errors is particularly relevant for modeling *adversarial noise*, where an adversary is allowed to decide which error operator to apply based on the specific codeword it acts upon. In Ref. (Ben-Aroya and Ta-Shma, 2009) it has shown that no QECC can perfectly correct controlled bit-flip errors, although a QECC of arbitrarily high dimension can approximately correct them (i.e., it can give back a codeword ‘close’ to the original one once an error acted on it).

B. Decoherence free subspaces

A suitable strategy to deal with completely correlated errors is represented by Noiseless Codes, also known as Decoherence Free Subspaces (DFS). This is a passive quantum error correction method where the key idea is that of avoiding decoherence by encoding quantum information into special subspaces that are protected from the interaction with the environment by virtue of some specific dynamical symmetry, see e.g. Lidar and Whaley, 2003.

To establish a direct link between QECCs and DFSs, one writes the evolution of the system Q density operator as

$$\Phi^{(n)}(\rho_Q) = \text{Tr}_B [U(\rho_Q \otimes \rho_B)U^\dagger] = \sum_{\mathbf{k}} A_{\mathbf{k}} \rho_Q A_{\mathbf{k}}^\dagger, \quad (52)$$

where $U = e^{-i(H_Q + H_B + H_{\text{int}})t}$ is the unitary evolution operator for the system-bath closed system and $A_{\mathbf{k}}$ the Kraus operators.

It turns out that a subspace \mathcal{H}_{DFS} of \mathcal{H} is a DFS if and only if all Kraus operators, when restricted to \mathcal{H}_{DFS} , are equal, up to a multiplicative constant, to a unitary transformation $U_Q^{(\text{DFS})}$. In this case, in a suitable basis,

the matrix representation of the Kraus operators is given by

$$A_{\mathbf{k}} = \begin{pmatrix} g_{\mathbf{k}} U_Q^{(\text{DFS})} & 0 \\ 0 & M_{\mathbf{k}} \end{pmatrix}, \quad (53)$$

where $M_{\mathbf{k}}$ is an arbitrary matrix that acts on $\mathcal{H}_{\text{DFS}}^\perp$ (with $\mathcal{H} = \mathcal{H}_{\text{DFS}} \oplus \mathcal{H}_{\text{DFS}}^\perp$) and may cause decoherence there. Equation (53) implies

$$A_{\mathbf{k}}^\dagger A_{\mathbf{l}} = \begin{pmatrix} \gamma_{\mathbf{kl}} I_C & 0 \\ 0 & M_{\mathbf{k}}^\dagger M_{\mathbf{l}} \end{pmatrix}, \quad (54)$$

where $\gamma_{\mathbf{kl}} = \bar{g}_{\mathbf{k}} g_{\mathbf{l}}$. Applying condition (43) to the present setting, it follows that DFS can be viewed as a special class of QECCs, where upon restriction to the code space $\mathcal{C} \equiv \mathcal{H}_{\text{DFS}}$, all recovery operators $R_{\mathbf{r}}$ are proportional to the inverse of the unitary $U_Q^{(\text{DFS})}$,

$$R_{\mathbf{r}} \propto \left(U_Q^{(\text{DFS})} \right)^\dagger. \quad (55)$$

It is worth noticing that in the DFSs case the matrix $\gamma_{\mathbf{kl}}$ has rank 1. Hence, a DFS is an example of maximally degenerated quantum error correcting code.

To show the effectiveness of DFS for completely correlated errors, let us consider a simple example $\Phi_{\mu=1}$ in (50) with $n = 2$ and $A_0 = \sigma_0$, $A_1 = \sigma_z$, $p_0 = 1 - p$, $p_1 = p$ so that

$$\begin{aligned} A_{00} &= \sqrt{p} \sigma_0 \otimes \sigma_0, & A_{01} &= 0, \\ A_{10} &= 0, & A_{11} &= \sqrt{(1-p)} \sigma_z \otimes \sigma_z. \end{aligned} \quad (56)$$

Clearly the DFS is given by $\text{span}\{|01\rangle, |10\rangle\}$ where one can safely encode a qubit

$$|0_L\rangle = |01\rangle, \quad |1_L\rangle = |10\rangle. \quad (57)$$

By extending this argument, one can say that in the case of completely correlated errors it is possible to exploit the invariance of a subspace to safely encode information.

In Ref. (Chiribella *et al.*, 2010), it has been provided a generalized quantum Hamming bound for nondegenerate codes, which depends on the rank of the Choi-Jamiołkowski operator representing the noise process and holds for any kind of noise. The original Hamming bound (45), which was formulated for the case of independent noise on the encoding systems is then recovered as a particular case. For completely correlated noise it has been shown how to exploit degeneracy to violate the generalized quantum Hamming bound and achieve perfect quantum error correction with fewer resources than those needed for non-degenerate codes. As an example consider the following channel

$$\begin{aligned} \Phi(\rho) &= p \rho + \sum_{i=1, \dots, n, j > i} \left(p_{X,ij} \sigma_x^{(i)} \sigma_x^{(j)} \rho \sigma_x^{(i)} \sigma_x^{(j)} \right. \\ &\quad \left. + p_{Y,ij} \sigma_y^{(i)} \sigma_y^{(j)} \rho \sigma_y^{(i)} \sigma_y^{(j)} + p_{Z,ij} \sigma_z^{(i)} \sigma_z^{(j)} \rho \sigma_z^{(i)} \sigma_z^{(j)} \right), \end{aligned} \quad (58)$$

where the input state is left unchanged with probability $p = 1 - \sum_{i=1, \dots, n, j > i} (p_{X,ij} + p_{Y,ij} + p_{Z,ij})$, while it undergoes Pauli operators σ_x , σ_y and σ_z on qubits i and j with probabilities $p_{X,ij}$, $p_{Y,ij}$ and $p_{Z,ij}$ respectively. By evaluating the rank of the map (58) one can see that the generalized quantum Hamming bound of Ref. (Chiribella *et al.*, 2010) reads in this case

$$2^k \left[1 + 3 \binom{n}{2} \right] \leq 2^n. \quad (59)$$

Then, by considering for instance $k = 1$ one gets $n = 7$ as smallest integer satisfying the bound. However, one can also construct codes with lower values for n . As matter of fact consider the following code

$$|0_L\rangle = |000\rangle, \quad |1_L\rangle = |111\rangle. \quad (60)$$

Now notice that these codewords are not affected by the action of σ_z on any pair of qubits. Consequently the action of σ_x on a pair of qubits is identical to the action of σ_y on the same pair of qubits. In other words the code is degenerate. Therefore, one has only to correct errors due to σ_x operators. This can be realized through a projective measurement onto the subspaces $S_{00} = \text{span}\{|000\rangle, |111\rangle\}$, $S_{01} = \text{span}\{|100\rangle, |011\rangle\}$, $S_{10} = \text{span}\{|010\rangle, |101\rangle\}$ and $S_{11} = \text{span}\{|001\rangle, |110\rangle\}$. If the measurement outcome is 00, no errors have affected the qubits; on the contrary, if the measurement outcome is 01, errors have affected qubits 2 and 3 and can be corrected by applying there σ_x ; if the measurement outcome is 10, errors have affected qubits 1 and 3 and can be corrected by applying there σ_x ; finally if the measurement outcome is 11, errors have affected qubits 1 and 2 and can be corrected by applying there σ_x . It is clear that this code violates the quantum Hamming bound (59) thanks to the invariance of the coding subspace under the action of pair of σ_z which allows for perfect error correction.

1. Threshold values for the degree of correlation

Having seen that DFSs are suitable to encode information in the presence of completely correlated errors, it is natural to expect that their performance decrease by reducing the degree of errors' correlation (Demkowicz-Dobrzanski, Kolenderski and Banaszek, 2007).

In Ref. (Cafaro and Mancini, 2010a) the model of (50) with $A_0 = \sigma_0$, $A_1 = \sigma_x$, $p_0 = 1 - p$ and $p_1 = p$ has been considered using a noiseless code of variable length (number of physical qubits) and fixed size (number of logical qubits). It has turned out that the entanglement fidelity (49), while obviously increasing with n , is also with μ and achieves the maximum allowed value of 1 for $\mu = 1$.

In Ref. (Cafaro and Mancini, 2011) the model of (51) with $A_0 = \sigma_0$, $A_1 = \sigma_x$, $p_0 = 1 - p$ and $p_1 = p$ has been considered using a noiseless code of variable length (number of physical qubits) and fixed size (number of

logical qubits). Also in this case the entanglement fidelity (49), while obviously increasing with n , is increasing with μ , achieving the maximum allowed value of 1 when $\mu = 1$.

These results are opposite to those reviewed in the previous Section about repetition codes. Then, Refs. (Cafaro and Mancini, 2011) and (Cafaro and Mancini, 2011) pointed out that there is a certain threshold value ($\mu^*(p)$) of μ depending on p , below which the repetition code works better than the noiseless one and above which the noiseless code works better than the repetition one. These results are summarized in Fig. 5.

Similar results were obtained in Ref. (D'Arrigo *et al.*, 2008), where it has been considered a Markovian correlated dephasing channel. In the system-environment Hamiltonian, the l -th carrier (qubit) interacts with the environment by means of a single Pauli operator, $\sigma_z^{(l)}$. Dephasing channels are thus characterized by the property that, when N qubits are sent through the channel, the states of a preferential orthonormal basis $\{|j\rangle \equiv |j_1, \dots, j_N\rangle, j_1, \dots, j_N = 0, 1\}$, with $\{|j_l\rangle\}$ eigenvectors of $\sigma_z^{(l)}$, are transmitted without errors. Of course superpositions of basis states may decohere, thus corrupting the transmission of quantum information. This model has been used to compare the performance of the standard three-qubit repetition code with a two-qubit noiseless code.

More generally, one may argue that, for the memory channel models (50), (51), there must be a threshold value $\mu^*(p)$ that allows us to select the best code between the standard and the noiseless ones.

C. Designing quantum codes for correlated errors

The specific features of error models can be used to design new quantum codes that better cope with correlated errors.

For instance, in (Novais and Baranger, 2006) it is studied the decoherence of a quantum computer in an environment which is inherently correlated in time and space. They first derive the nonunitary time evolution of the computer and environment in the presence of a stabilizer error correcting code, providing a general way to quantify decoherence. The general theory is then applied to the spin-boson model. The results demonstrate that effects of long-range correlations can be systematically reduced by small changes in the error correction codes.

More generally, one can summarize the following developments.

1. Concatenated codes

The results at the end of the previous Subsection suggest that it may be convenient to concatenate decoherence-free subspaces with standard quantum error correcting codes in order to achieve higher entanglement fidelity values in both low and high correlations regimes.

This kind of concatenation was first introduced in Ref. (Lidar, Bacon and Whaley, 1999), however in Ref. (Clemens, Siddiqui and Gea-Banacloche, 2004) and subsequently in Ref. (Cafaro and Mancini, 2011) it was investigated in the context of memory channels.

For the sake of simplicity, let us illustrate the case of two layers of concatenation and consider single qubit encoding. Assume the inner code (first layer) is a $[[n_1, k_1, d_1]]$ stabilizer code \mathcal{C}_1 with generators $G_1 = \{g_i^1 : i = 1, \dots, n_1 - k_1\}$, and the outer code (second layer) is a $[[n_2, 1, d_2]]$ stabilizer code \mathcal{C}_2 with generators $G_2 = \{g_j^2 : j = 1, \dots, n_2 - 1\}$. The concatenated code $\mathcal{C} = \mathcal{C}_1 \circ \mathcal{C}_2$ maps k_1 qubits into $n = n_1 n_2$ qubits, with code construction parsing the n qubits into n_1 blocks $B(b)$ ($b = 1, \dots, n_1$) each containing n_2 qubits. In other words, given a codeword $|c_{\text{in}}\rangle$ for the inner code \mathcal{C}_1 ,

$$|c_{\text{in}}\rangle = \sum_{j=0}^{k_1-1} \alpha_j |\phi_j\rangle, \quad (61)$$

where $\{|\phi_j\rangle\}$ are basis vectors for \mathcal{C}_1 , the concatenated code \mathcal{C} is constructed as follows. For any codeword $|c_{\text{out}}\rangle$ for the outer code \mathcal{C}_2 ,

$$|c_{\text{out}}\rangle = \sum_{i_1, \dots, i_{n_2}} \alpha_{i_1, \dots, i_{n_2}} |i_1 \dots i_{n_2}\rangle, \quad (62)$$

with $|i_1 \dots i_{n_2}\rangle = |i_1\rangle \otimes \dots \otimes |i_{n_2}\rangle$, replace each basis vector $|i_j\rangle$ with $i_j = 0, \dots, k_1 - 1$ for $j = 1, \dots, n_2$ by a basis vector $|\phi_{i_j}\rangle$ in \mathcal{C}_1 , that is

$$|c_{\text{conc}}\rangle := \sum_{i_1, \dots, i_{n_2}} \alpha_{i_1, \dots, i_{n_2}} |\phi_{i_1}\rangle \otimes \dots \otimes |\phi_{i_{n_2}}\rangle. \quad (63)$$

Further details on the construction of the stabilizer generators of \mathcal{C} can be found in (Gaitan, 2008). As a final remark, one simply points out that the above mentioned construction produces a $[[n_1 n_2, k_1, d]]$ code with $d \geq d_1 d_2$.

As an illustrative example, consider to encode one logical qubit with a concatenated subspace obtained by combining the decoherence free subspace (for a completely correlated bit flips)

$$|0_L\rangle = |+-\rangle, \quad |1_L\rangle = |-+\rangle, \quad (64)$$

(inner code, $\mathcal{C}_{DFS} = \mathcal{C}_{\text{inner}}$) with the three-qubit bit repetition code in (60) (outer code, $\mathcal{C}_{\text{bit}} = \mathcal{C}_{\text{outer}}$). One obtains that the codewords of the concatenated code $\mathcal{C} = \mathcal{C}_{DFS} \circ \mathcal{C}_{\text{bit}}$ are given by,

$$\begin{aligned} |0_L\rangle &= \frac{1}{2} (|000000\rangle - |000111\rangle + |111000\rangle - |111111\rangle), \\ |1_L\rangle &= \frac{1}{2} (|000000\rangle + |000111\rangle - |111000\rangle - |111111\rangle). \end{aligned}$$

The entanglement fidelity for the concatenation of a repetition code and a noiseless code for the model of (50)

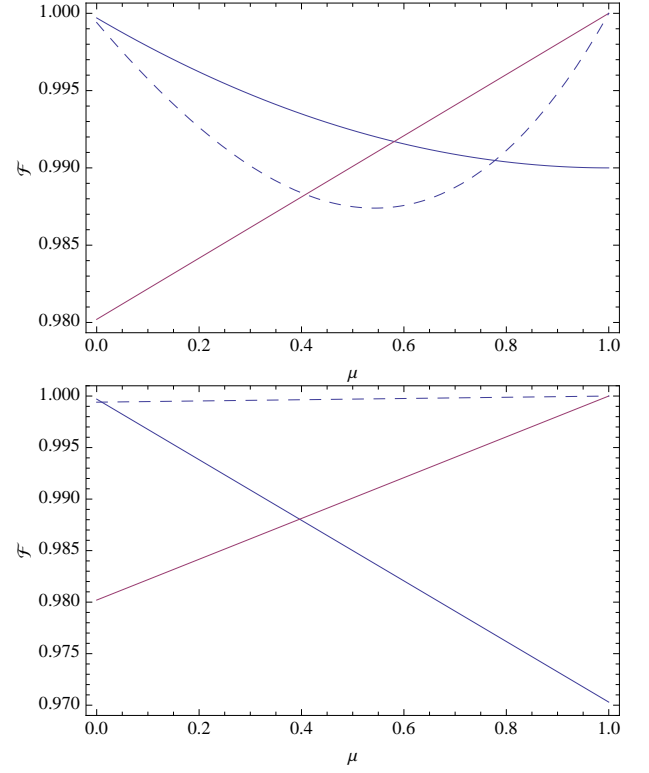


FIG. 5 Entanglement fidelity for the model (50) top figure and (51) bottom figure with $p = 10^{-2}$. Blue solid lines correspond to standard code. Red solid lines correspond to noiseless code. Dashed lines correspond to their concatenation.

with $A_0 = \sigma_0$, $A_1 = \sigma_x$, $p_0 = 1 - p$ and $p_1 = p$ are reported in Fig. 5. It turns out that the concatenated code does not work well for partially correlated errors. It is always better to use either the outer or the inner code alone depending on whether one is below or above the threshold value $\mu^*(p)$.

The entanglement fidelity for the concatenation of repetition code and noiseless code for the model of (51) with $A_0 = \sigma_0$, $A_1 = \sigma_x$, $p_0 = 1 - p$ and $p_1 = p$ are shown in Fig. 5 too. Here the concatenated code works perfectly almost everywhere. Hence, one may argue that for the model of (51) the concatenation trick results particularly advantageous in the presence of partially correlated errors.

2. Burst errors codes

In communication channels noise disturbances over long time periods lead to temporally continuous errors, as opposed to noise disturbances over short time periods leading to random single errors. Analogously, stored data can be defective over spatially length scales greater than a single bit. Such errors are called burst errors. They are well studied in the classical framework where corresponding error correcting codes have been developed (Peterson and Weldon, 1972). In (Vatan, Roychowdh-

hury and Anantram, 1997), quantum analog of burst-error correcting codes have been considered. Hamming and Gilbert-Varshamov type bounds have been derived showing that these codes are more efficient than codes protecting against random errors. In fact, to protect against burst errors of width b (that is, errors occurring on a number b of consecutive qubits with b a fixed constant), it is enough to map $n - \log n - O(b)$ qubits to n qubits, while in the case of t random errors at least $n - t \log n$ qubits should be mapped to n qubits. Based on binary cyclic codes explicit constructions have been presented of these almost optimal quantum code for bursts of width not bigger than 4 and for specific values of n . It remains open the problem of how to generalize these constructions.

Quite generally, one must increase the redundancy of the codeword with increasing length of the burst. On the other hand, if details of the decoherence mechanism of qubits are known, then it might be possible to design a more efficient error-correction method than the conventional scheme. In (Kawataba, 2000), it is proposed a method for constructing quantum burst-error-correcting codes from known quantum error-correcting codes. This method is based on the interleaving technique. By using this method, the quantum codewords can be distributed amongst the qubit stream so that consecutive words are never next to each other. On de-interleaving they are returned to their original positions so that any errors that have occurred become widespread. This ensures that any burst (long) errors now appear as random (short) errors.

3. Convolutional codes

It is possible (Chau, 1998, 1999; Ollivier and Tillich, 2003, 2004) to extend the notion of stabilizer codes to codes which allow for an overlap between the individual steps of the encoding operation. It can be noticed that their encoding operations can be described as quantum memory channels (Gütschow, 2010), due to the fact that some of the output qubits of the n th encoding step will be used as inputs in the $(n + 1)$ th step of the encoding, thus the blocks overlap. These qubits correspond to the memory system of the memory channel describing the encoding. The encoding operation is the same in every step (neglecting initialization and finalization), thus every step is described by the same channel. From the classical coding theory, codes with these properties are called convolutional codes.

In particular, convolutional stabilizer codes correspond to Clifford memory channels (Gütschow *et al.*, 2010). Looking at the encoding operation as a quantum channel, it is convenient to take the encoding of one block as one use of the channel. As one passes on qubits from one step of the encoding to the next, one deals with a memory channel. All encoding steps are the same, so our channel can be concatenated and used for arbitrary many steps of the encoding procedure. Stated more pre-

cisely, for every $[[n, k, m]]$ convolutional stabilizer code one can find an encoding operation which is described by a the concatenation of a reversible Clifford memory channel. The channel has n input and output qubits and uses m qubits of memory. In this framework the memory system appears as a resource for the implementation of the correcting code.

Unfortunately convolutional codes also carry disadvantages. Because information is transmitted from one block to the next, errors can spread as well. Depending on the encoding algorithm catastrophic errors can occur. Catastrophic errors happen during the process of encoding (or transmission or decoding) and affect only a few qubits at first. During the encoding (or decoding) they spread without bound. These errors can only be corrected with operations of unbounded support (i.e., they act on an unbounded number of qubits). If one considers transmissions of finite length, the support of the operation is only bounded by the length of the transmission. There is one exception of this behavior. The support of a catastrophic errors can also stay finite but the localization area moves in an unbounded manner, i.e., the error will always affect the qubits at the end of the transmission. In both cases the error can only be corrected at the end of the transmission. This implies that the decoding process can start only after the end of the transmission. Therefore the delay between receiving the first qubit and decoding it is as long as the transmission itself. Furthermore the memory requirement on the decoder side is of the order of the length of the encoded message. Consequently these errors have to be avoided by employing non-catastrophic convolutional codes (Grassl and Roetteler, 2006). Non-catastrophic codes with minimal dimension of the memory system can be devised as well (Houshmand, Hosseini-Khayat and Wilde, 2011; Wilde *et al.*, 2011). Moreover, the dimension of the memory system of non-catastrophic codes can be further optimized by allowing the assistance of pre-shared entanglement (Wilde and Brun, 2009, 2010a,b). The link between Clifford memory channels and convolutional stabilizer codes allows to prove that a convolutional encoder is non-catastrophic if and only if the corresponding memory channel representing is strictly forgetful (Gütschow, 2010).

Not only convolutional codes can be described by memory channels, but also memory channels could benefit from convolutional codes. In fact, when sending information across correlated channels, data put into one channel may influence the output of all further invocations. Therefore, a decoding circuit needs to gather information from the output of several channel uses in order to decode any given bit. Block codes are ill-adapted to this task, as one expects a considerable loss of information to occur at every block boundary. Convolutional codes could be much better suited for this particular situation. Lacking the boundaries between code blocks, convolutional codes exhibit the same "continuous structure" as channels with memories. Convolutional decoding circuits

continuously gather output from the channel, outputting decoded qubits as soon as sufficient data has become available.

VI. CODING THEOREMS AND CAPACITIES

Here, one discusses more in detail and quantify the capability of a quantum channel Φ to transfer classical and quantum information encoded on quantum systems. This will lead to the notions of capacities of a quantum channel, which are the quantum counterparts of the Shannon capacity (Shannon, 1948).

In the context of classical communication theory, the latter was defined as the optimal rate of reliable information transmission, that is the maximum rate at which information can be sent through a channel with asymptotically vanishing error probability, in the limit of long messages, and by using proper encoding and decoding procedures (asymptotically eliminating all errors). Shannon proved that this quantity is expressed in terms of a maximization of an entropic quantity (mutual information) between the two parties.

Being non constructive, this proof does not provide an explicit recipe for the optimal code. Roughly speaking, the codewords of the optimal code correspond to distinguishable output states, in such a way one can efficiently recover the input with an arbitrarily small error.

The generalization of this concept to the quantum setting leads to the definition a family of channel capacities, depending on whether classical or quantum information has to be transmitted, and whether additional resources, as pre-shared entanglement, are exploited. Indeed, the capacities of a quantum channel are operationally defined as the optimal communication rates through the given quantum channel. Then, coding theorems prove their expression in terms of entropic quantities (see Appendix B).

A. Operational definitions and memoryless setting

The classical (quantum) capacity of a quantum channel is then the maximum rate at which classical (quantum) information, encoded on a quantum system, can be reliably sent from the sender to the receiver through the given quantum channel. This rate basically is the maximum ratio among the number of bits (qubits) transmitted and the “redundancy” employed into the code. In mathematical terms, the classical capacity $C(\Phi)$ of a noisy quantum channel Φ can be defined as follows (Bennett and Shor, 1998)

$$C(\Phi) := \lim_{\epsilon \rightarrow 0} \limsup_{k \rightarrow \infty} \left\{ \frac{k}{n} : \exists_{n, \Phi_E^{(k \rightarrow n)}, \Phi_D^{(n \rightarrow k)}} \forall_{|\psi\rangle \in \{|0\rangle, |1\rangle\}^{\otimes k}} \right.$$

such that

$$F(|\psi\rangle\langle\psi|, \Phi_D^{(n \rightarrow k)} \circ \Phi^{\otimes n} \circ \Phi_E^{(k \rightarrow n)}(|\psi\rangle\langle\psi|)) > 1 - \epsilon \Big\}, \quad (65)$$

with $F(\rho_1, \rho_2)$ being the quantum fidelity between the states ρ_1 and ρ_2 as defined in Eq. (A3) in Appendix A, Φ_E being an encoding map from n qubits to m inputs of the channel Φ , and Φ_D a decoding map from m channel outputs and n qubits. Let us point out that the input state $|\psi\rangle$ is simply a tensor product of states $|0\rangle$ and $|1\rangle$, hence encoding classical information. Indeed, the definition of $C(\Phi)$ is related to the transmission of classical information encoded in quantum states, hence no superposition states are considered here. In other words, $C(\Phi)$ is defined as the optimal rate (i.e., transmitted bits per channel use) at which the sender can send a tensor product state, i.e. $|\psi\rangle \in \{|0\rangle, |1\rangle\}^{\otimes n}$, of n qubits, for arbitrarily large n , to the receiver, who is able to recover it with fidelity greater than $1 - \epsilon$, with arbitrarily small ϵ , after encoding, transmission, and decoding procedures.

Another quantity of interest is the *one-shot* capacity, or *product-state* capacity, denoted $C_1(\Phi)$, which is defined as in (65) with additional assumption that the employed codes are made of separable states, that is, $\Phi_E^{(k \rightarrow n)}(|\psi\rangle\langle\psi|)$ is a separable state for all $|\psi\rangle \in \{|0\rangle, |1\rangle\}^{\otimes k}$.

Similarly, the quantum capacity $Q(\Phi)$ is defined as (Bennett and Shor, 1998)

$$Q(\Phi) := \lim_{\epsilon \rightarrow 0} \limsup_{k \rightarrow \infty} \left\{ \frac{k}{n} : \exists_{n, \Phi_E^{(k \rightarrow n)}, \Phi_D^{(n \rightarrow k)}} \forall_{|\psi\rangle \in \mathcal{H}^{\otimes k}} \right.$$

such that

$$F(|\psi\rangle\langle\psi|, \Phi_D^{(n \rightarrow k)} \circ \Phi^{\otimes n} \circ \Phi_E^{(k \rightarrow n)}(|\psi\rangle\langle\psi|)) > 1 - \epsilon \Big\}, \quad (66)$$

where \mathcal{H} is the two-dimensional Hilbert space of a two-level quantum system. Notice that the only difference between the definitions of C and Q is that now the information is quantum and can be encoded also on superpositions of k qubits. Since a tensor product state is always a specific case of a state living in \mathcal{H} , one obtains the first of a series of inequalities between channel capacities, i.e. $Q(\Phi) \leq C(\Phi)$. This crucial difference between C and Q allows one to exploit some peculiar properties of quantum states, as entanglement, in order to build up optimal encoding inputs. Analogously, joint quantum measurements of the output states have to be taken into account in the choice of the optimal decoding strategies.

The *entanglement-assisted classical capacity*, denoted $C_{ea}(\Phi)$, is defined as the maximum rate of reliable transmission of classical information when the sender and the receiver have at their disposal an unbounded number of pre-shared maximally entangled states. Coding theorems prove that the entanglement-assisted classical capacity of a memoryless channel is given in terms of the quantum mutual information [defined in Eq. (B4) in Appendix B] (Bennett *et al.*, 1999, 2002):

$$C_{ea}(\Phi) = \max_{\rho} \mathcal{I}(\rho, \Phi), \quad (67)$$

where the maximization is over any input ρ .

Concerning the (unassisted) classical capacity, let us first focus on the one-shot classical capacity, whose expression is given by (Holevo, 1998; Schumacher and Westmoreland, 1997)

$$C_1(\Phi) = \max_{\{p_i, \rho_i\}} \left[S\left(\sum_i p_i \Phi[\rho_i]\right) - \sum_i p_i S(\Phi[\rho_i]) \right], \quad (68)$$

where the maximum is taken over all probability distributions $\{p_i\}$ and collections of density operators $\{\rho_i\}$ (possibly satisfying some additional input constraints). The quantity in bracket in Eq. (68) is known as *Holevo information* and denoted as $\chi(\Phi, \{p_i, \rho_i\})$. The expression for the classical capacity of the quantum channel is then obtained by allowing encoding blocks of increasing length n (Holevo, 1998; Schumacher and Westmoreland, 1997), i.e.,

$$C(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} C_1(\Phi^{\otimes n}). \quad (69)$$

Since an additional resource (shared entanglement) is required for the definition of $C_{ea}(\Phi)$, the following inequality holds $C(\Phi) \leq C_{ea}(\Phi)$. Finally, let us point out that the additivity conjecture, $C(\Phi) = C_1(\Phi)$, i.e. entangled input states do not increase the classical capacity, was believed to be true, allowing a practical computation of this quantity. However, recently, it has been shown that this conjecture is not always true, by constructing a counter-example where the one-shot capacity C_1 is superadditive, i.e. $C_1(\Phi \otimes \Phi) > 2C_1(\Phi)$ with Φ being a random-unitary quantum channel (Hastings, 2009). Notwithstanding, one may safely say that additivity holds true for a large variety of channels (Amosov and Mancini, 2009; Hiroshima, 2006). Note also that the additivity counter-example does not necessarily imply that $C(\Phi_1 \otimes \Phi_2) > C(\Phi_1) + C(\Phi_2)$ since an additional regularization is needed in computing C as in Eq. (69). These results leave still open the problem of channel capacity, i.e. the full optimization over all entangled input states is generally required to calculate the maximum capacity C .

The expression for the quantum capacity $Q(\Phi)$ of a memoryless channel Φ is provided by (Barnum, Nielsen and Schumacher, 1998; Devetak, 2005; Lloyd, 1997; Shor, 2002)

$$Q(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} Q_1(\Phi)^{(n)}, \quad (70)$$

where

$$Q_1(\Phi)^{(n)} = \max_{\rho} J(\rho, \Phi^{\otimes n}), \quad (71)$$

where the maximization is performed over all input states for n uses of the quantum channel and $J(\rho, \Phi^{\otimes n})$ is the *coherent information* of the map $\Phi^{\otimes n}$, as defined in Appendix B. The maximum of J may be strictly superadditive for parallel channels, hence the quantum

capacity may be greater than its *one-shot* expression, i.e. $Q(\Phi) > Q_1(\Phi)$, in contrast to the case of the entanglement-assisted classical capacity C_{ea} which is always additive. For this reason, the regularization for n which tends to infinite (i.e., $\lim_{n \rightarrow \infty}$) is necessary. On top of that, the coherent information is not, in general, a concave function, therefore there might be local maxima which are not global ones and this makes the calculation of Q one of the most difficult task in quantum communication theory. Notice that for degradable channels, defined in Sec. II, the coherent information can be proved to be additive and then the regularization for n is not necessary, i.e. $Q(\Phi) = Q_1(\Phi) = \max_{\rho} J(\rho, \Phi)$. On the other side, for anti-degradable channels Φ one always has $Q(\Phi) = 0$. The latter holds also for the class of PPT maps (see Section II.A). The three channel capacities, defined so far, satisfy the following inequality

$$Q(\Phi) \leq C(\Phi) \leq C_{ea}(\Phi). \quad (72)$$

Finally, recently an example of a new interesting phenomenon, called *superactivation*, has been provided (Smith and Yard, 2010). In particular, it has been shown that it is possible to find channels Φ_1 and Φ_2 with zero capacity, i.e. $C(\Phi_1) = Q(\Phi_1) = C(\Phi_2) = Q(\Phi_2) = 0$, for which, by joining the two communication lines Φ_1 and Φ_2 in $\Phi_1 \otimes \Phi_2$, it becomes possible to transmit quantum information, i.e. $Q(\Phi_1 \otimes \Phi_2) > 0$. Specifically, Φ_1 and Φ_2 are given by an anti-degradable channel and a PPT one, – see (Brandão, Oppenheim and Strelchuk, 2012) for other examples in terms of *depolarizing* maps and for a more general construction.

B. Entropic upper bounds: memory setting

The operational definitions of channel capacities, introduced above for the memoryless setting, apply also straightforwardly to memory quantum channels. Indeed, they are quite generally related to the optimal classical and quantum information transmission rates between two parties, no matter how complex is the internal structure of the channel. However, the memory setting is often more complicate than the memoryless one since additional channel features have to be specified in order to discuss the concept of communication capacity. For instance, in the case of non-anticipatory channels in Sec. III.B.1, the map can be described in terms of an additional memory system M , and then the initial and final memory states have to be characterized before defining the channel capacities. In particular, one has to distinguish different setups where Alice, Bob and/or Eve (third party) may control (or use for the encoding/decoding procedures) the initial/final states of the memory system (Kretschmann and Werner, 2004). Hence, this leads to different definitions and bounds for the memory channel capacities. In fact, for a channel Φ one can at least define four classical capacities $C_{AB}(\Phi)$, $C_{AE}(\Phi)$, $C_{EB, \mu}(\Phi)$, $C_{EE, \mu}(\Phi)$, with the first (second) index representing the

party controlling the initial (final) memory state and μ being the Eve's choice for the initial state of M (when considered). The same classification holds for the quantum capacities $Q_{AB}(\Phi)$, $Q_{AE}(\Phi)$, $Q_{EB,\mu}(\Phi)$, $Q_{EE,\mu}(\Phi)$. Moreover, it may be generalized to setups where additional resources, as shared entanglement and classical side (forward and/or backward) communication, do assist the memory noisy transmission.

1. Finite-memory channels

The first attempts to extend the notion of classical and quantum capacity to memory channels were shown in Ref. (Bowen and Mancini, 2004), where some bounds in terms of mutual and coherent information were provided, independently from the internal evolution of the memory state. In particular, they considered finite-memory channels, i.e. maps with memory of finite dimension (described in Sec. III.B.2), exploiting the Holevo quantity as an upper bound for the classical capacity $C(\Phi)$, i.e. for a given initial memory state Ω_M ,

$$C(\Phi) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\{p_i, \rho_i\}} \chi(\Phi^{(n)}, \{p_i, \rho_i\}) \quad (73)$$

where $\chi(\dots)$ is the Holevo quantity defined in Eq. (68), $\Phi^{(n)}$ as in Eq. (19), and ρ_i denotes a set of Alice states (shared entangled states between Alice and Bob) for unassisted (entanglement-assisted) communication. Unlikely in the memoryless case, here this upper bound is not always achievable. Indeed, the coding for each channel cannot be divided into smaller blocks because the memory state can be entangled over several blocks. However, this bound turns out to be achievable for a class of Markovian channels, i.e. maps represented by unitary Kraus operators and steady-state probabilities as initial error distributions, as defined in Sec. IV.C (Bowen and Mancini, 2004). The intuitive explanation for this bound achievement is that the memory plays a small role on the asymptotic behaviour of the channel. In other words, for Markovian memory channels the capacity is only affected by the loss information to the external environment, asymptotically neglecting its loss into the memory system. Analogously, one can derive a bound for the quantum capacity in terms of the coherent information, i.e.

$$Q(\Phi) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\rho} J(\rho, \Phi^{(n)}) \quad (74)$$

2. Perfect memory channels

These maps can be described in terms of Kraus operators whose number of elements is sub-exponentially growing in the number of channel uses n - see more detail in Sec. IV.B. In other terms, the size of the environment is not large enough to include the information

sent from Alice to Bob, which is exponentially increasing in n , and so asymptotically the lost of information into the environment is negligible. This intuitively explains that perfect memory channels are asymptotically noiseless and have maximal capacities. Specifically, one can verify that the following theorem holds (Giovannetti, Burgarth and Mancini, 2009; Kretschmann and Werner, 2005):

Theorem 1 *Let \mathcal{L} be a multi-use perfect memory channel as in (17) and let $\{d_E^{(n)} : n \in \mathbb{N}\}$ be the sequence satisfying Eq. (29). Then for sufficiently large n there exists a zero-error classical code \mathcal{C} of size*

$$|\mathcal{C}| \geq \frac{(d_Q)^n}{(d_E^{(n)})^2}, \quad (75)$$

with a rate $R_{\mathcal{C}} \geq \log d_Q - \frac{2}{n} \log d_E^{(n)}$ that for $n \rightarrow \infty$ converges to the optimal value $\log d_Q$. Here, $d_Q^{(n)}$ is the size of the system carriers, i.e. $d_Q^{(n)} = \dim \mathcal{H}_Q^{(n)}$.

Analogously, for sufficiently large n there exists a zero-error quantum error correcting code \mathcal{Q} of size

$$|\mathcal{Q}| \geq \frac{(d_Q)^n}{(d_E^{(n)})^4 + (d_E^{(n)})^2} \quad (76)$$

with rate $R_{\mathcal{Q}} \geq \log d_Q - \frac{1}{n} \log[(d_E^{(n)})^4 + (d_E^{(n)})^2]$ which, again, for $n \rightarrow \infty$ converges to the optimal value $\log d_Q$.

3. Non-anticipatory memory channels

For generic non-anticipatory (n -fold concatenated) memory quantum channels $\Phi^{(n)}$ as in Eq. (17), the classical capacities are upper bounded as (Kretschmann and Werner, 2005):

$$C_{EE,\mu}(\Phi) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \max_{\{p_i, \rho_i\}} \chi(\text{tr}_M \circ \Phi^{(n)}, \{p_i, \mu \otimes \rho_i\}) . \quad (77)$$

Analogously, one finds the following upper bounds for the corresponding quantum capacities, i.e.

$$Q_{EE,\mu}(\Phi) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \max_{\rho} J(\text{tr}_M \circ \Phi^{(n)}, \mu \otimes \rho) . \quad (78)$$

Similar expressions are obtained in the other three configurations discussed in Sec. VI.B. The proof of this theorem comes from similar bounds for the memoryless settings. However, in this context coding theorems seem to be much harder to be proved, and these entropic bounds for classical and quantum information transmission rates are demonstrated to be achievable for a specific class of memory channels, called forgetful ones, as further discussed below.

4. Forgetful channels

These communication channels are characterized by the property that the effect of the initial memory state becomes negligible with time, i.e. memory effects die away exponentially fast, as discussed in Sec. IV.D. This feature allows one to prove that the upper bounds in Sec. VI.B.3 can be actually asymptotically achieved (Kretschmann and Werner, 2005). This important result can be demonstrated reducing the problem into the case of memoryless channel by means of a double-blocking procedure. In other words, one can define the *memory depth* of a forgetful channel Φ as the smallest integer $m \in \mathbb{N}$ such that

$$\Phi_m = (\mathcal{P} \otimes \text{id}_Q^{\otimes m}) \circ \Phi_m, \quad (79)$$

with \mathcal{P} being the completely *depolarizing* channel, and id_Q being the identity map acting on $\mathfrak{S}(\mathcal{H}_Q)$. If a finite m does exist, these maps are called *strictly forgetful*. Hence, it is possible to group the channels $\Phi^{(n)}$ into blocks of length $m + l$, encoding the input in the l channels and ignoring the intermediate m ones. The asymptotic behaviour of the channels is obtained when $l \rightarrow \infty$. Furthermore, restricting the inputs to product states of block length $m + l$ and exploiting the strict forgetfulness property, the output state factorizes, and the whole map is reduced to a memoryless channel on the larger input space $\mathcal{H}_Q^{\otimes l+m}$. It allows to extend Holevo and Devetak coding theorems for memoryless channels to forgetful memory channels. Note that the double-block strategy can be applied even if the channel is not strictly forgetful. Therefore, the memoryless expressions for classical and quantum channel capacities in Eqs. (69) and (70) can be applied also in the memory setting for forgetful maps, and the entropic upper bounds in Eqs. (77) and (78) are exactly achieved.

Forgetful channels have been proven to constitute a dense set with the topology induced by the norm of complete boundness (Kretschmann and Werner, 2005). That implies that any non-forgetful channel can be approximated by a forgetful one. Notwithstanding, their capacities may be different. An example is given by in the context of Markovian-correlated channels. A long term-memory channel, Sec. IV.E, can be approximated by a forgetful Markovian-correlated channel, Sec. IV.C. However, according to the coding theorem for long-term memory channel reviewed in the following section, the capacity of the latter does not approximate the capacity of the former.

5. Long-term memory channels

To provide coding theorems for channels with long-term memory (i.e., channels which are “not forgetful”) is more complicate. A working class of this kind of quantum channels has been identified in (Datta and Dorlas, 2007), where convex combinations of memoryless channels were

considered (see Sec. IV.E). For a channel Φ in this class, $\Phi^{(n)} : \mathfrak{S}(\mathcal{H}^{\otimes n}) \rightarrow \mathfrak{S}(\mathcal{K}^{\otimes n})$ and the action of $\Phi^{(n)}$ on any state $\rho^{(n)} \in \mathfrak{S}(\mathcal{H}^{\otimes n})$ is given as follows

$$\Phi^{(n)}(\rho^{(n)}) = \sum_{i=1}^M p_i \Phi_i^{\otimes n}(\rho^{(n)}) \quad (80)$$

where $\Phi_i : \mathfrak{S}(\mathcal{H}) \rightarrow \mathfrak{S}(\mathcal{K})$ ($i = 1, \dots, M$) are completely positive, trace preserving (CPT) maps and $p_i > 0$, $\sum_{i=1}^M p_i = 1$. Here, \mathcal{H} and \mathcal{K} denote Hilbert spaces. A classical version of this channel was introduced by Jacobs (Jacobs, 1962) and studied further by Ahlswede (Ahlswede, 1968).

The following coding theorem for quantum channels (80) has been derived in (Datta and Dorlas, 2007).

Theorem 2 *The one-shot capacity of a channel Φ , with long-term memory, defined through (80), is given by*

$$C_1(\Phi) = \sup_{\{p_j, \rho_j\}} \left[\min_i \chi(\Phi_i, \{p_j, \rho_j\}) \right], \quad (81)$$

The supremum is taken over all finite ensembles of states $\rho_j \in \mathfrak{B}(\mathcal{H})$ with probabilities p_j .

This theorem has been proved by employing a quantum version of Feinstein’s Fundamental Lemma (Feinstein, 1954; Jacobs, 1957) and a generalization of Helstrom’s theorem (Helstrom, 1976). For a quantum memoryless channel, the method yields an alternative proof of the Holevo-Schumacher-Westmoreland (HSW) theorem (Holevo, 1998; Schumacher and Westmoreland, 1997) similar to the proof in (Winter, 1999).

The one-shot capacity can be generalized to give the classical capacity of the channel Φ in the usual manner, that is, by considering inputs which are product states over uses of blocks of n channels, but which may be entangled across different uses within the same block. The classical capacity $C(\Phi)$ is obtained in the limit $n \rightarrow \infty$ and is given by

$$C(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} C_1(\Phi^{(n)}). \quad (82)$$

These results have been generalized to the case of a continuous set of channels by Bjelaković and Boche, 2009.

Then, a coding theorem for communication assisted by entanglement in the class of channels specified by (80) has been derived as well (Datta, Suhov and Dorlas, 2008).

Theorem 3 *The entanglement-assisted classical capacity of a channel Φ , with long-term memory, defined through (80), is given by*

$$C_{ea}(\Phi) = \max_{\rho \in \mathfrak{S}(\mathcal{H}_A)} \left[\min_i \mathcal{I}(\rho, \Phi_i) \right], \quad (83)$$

where $\mathcal{I}(\rho, \Phi_i)$ is the quantum mutual information (B4).

The proof of this theorem, makes use of the expression for the product state capacity (81).

The above theorems can be extended to quantum channels with arbitrary Markovian correlated noise as shown in (Datta and Dorlas, 2009). Finally, (Bjelaković, Boche and Nötzel, 2009) provided a formally analogous expression for the quantum capacity of long-term memory channels.

6. Ergodic cq-channels with decaying input memory

For ergodic cq-channel with decaying input memory a coding theorem has been derived in (Bjelaković and Boche, 2008).

Theorem 4 *Let $W : A^{\mathbb{Z}} \times \mathcal{B}^{\mathbb{Z}} \rightarrow \mathbb{C}$ be a stationary ergodic decaying input memory cq-channel, then the classical capacity is given by*

$$C(W) = \sup_{p \text{ stationary ergodic}} i(p, W), \quad (84)$$

where

$$i(p, W) := \lim_{n \rightarrow \infty} \frac{1}{n} (S(\rho_p^n) + S(\rho_W^n) - S(\rho_{p,W}^n)), \quad (85)$$

with

$$\rho_p^n = \sum_{x^n \in A^n} p^n(x^n) |x^n\rangle\langle x^n|, \quad (86)$$

$$\rho_W^n = \sum_{x^n \in A^n} p^n(x^n) \rho_{x^n}, \quad (87)$$

$$\rho_{p,W}^n = \sum_{x^n \in A^n} p^n(x^n) |x^n\rangle\langle x^n| \otimes \rho_{x^n}. \quad (88)$$

Here ρ_{x^n} denotes the density operator of the output state $W^n(x^n, \cdot)$, $x^n \in A^n$ and $|x^n\rangle = e_{x_1} \otimes \dots \otimes e_{x_n}$ for some orthonormal basis $\{e_i\}_{i=1}^{|A|}$ of $\mathbb{C}^{|A|}$.

The sup in Eq.(84) is calculated over all stationary ergodic probability measures p on $A^{\mathbb{Z}}$. That is, consider a shift $T : A^{\mathbb{Z}} \rightarrow A^{\mathbb{Z}}$ of double infinite sequences of A , then p is stationary if $p(Ta) = p(a)$ for all $a \in A^{\mathbb{Z}}$. Moreover, it is ergodic if for all $a \in A^{\mathbb{Z}}$ such that $Ta = a$ it is $p(a) = 0$ or 1.

The above theorem results as an extension of coding theorem for input memoryless cq-channel whose proof combines Wolfowitz's code construction (Wolfowitz, 1957) and a version of the Feinstein's lemma (Blackwell, Breiman and Thomasian, 1958) based on the notion of the joint input output probability distribution.

VII. SOLVABLE MODELS OF MEMORY CHANNELS

Up to date, only few models of memory quantum channels have been fully solved in terms of capacities. One is the dephasing channel (in the discrete variable setting), and the other is the lossy bosonic channel (in the continuous variable setting), with different types of correlations

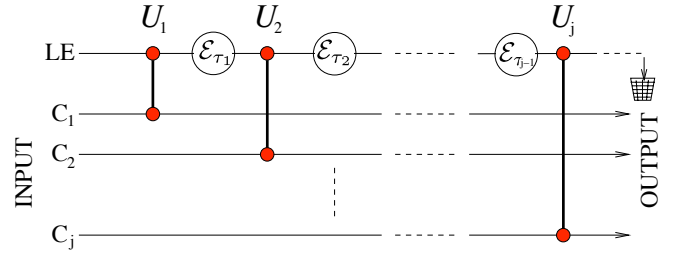


FIG. 6 Circuit representation of Eq. (91). The local environment LE interacts through the unitary couplings U_j with one carriers at a time. Between two consecutive interactions with the carriers instead LE undergoes the dissipative evolution described by the transformations \mathcal{E}_{τ_j} .

A. Discrete memory channels

The correlated dephasing channel was first discussed in (Giovannetti, 2005) as an example of dynamical memory model. Such a model consists in a communication line where messages are encoded into identical carriers q_1, q_2, \dots propagating through a medium E that separates the sender (Alice) from the receiver (Bob).

The carriers are organized in a time-ordered sequence $s = \{\tau_1, \tau_2, \dots\}$ with $\tau_j > 0$ being the time interval between the instants t_{j+1} and t_j at which q_{j+1} and q_j enter E respectively. The effective transit time is assumed to be constant and much shorter than the intervals τ_j .

Two distinct components of the medium E are identified: a finite dimensional Local Environment (LE) component which is directly coupled with the carriers through unitaries U_j , and a huge Reservoir (R) component which is coupled with LE but not with the carriers. The action of R on LE is supposed to induce a dissipative dynamics which transforms any initial states of LE into a stationary configuration σ_0 , with τ_E being the characteristic time of the process. This is equivalent to introducing a one-parameter family $\{\mathcal{E}_\tau\}_{\tau \geq 0}$ of CPTP maps which, given σ the initial state of LE at some time t_0 , represents its evolution at time $t_0 + \tau$ with the density matrix $\mathcal{E}_\tau(\sigma)$. In this formalism \mathcal{E}_0 coincides with identity map on \mathcal{H}_{LE} . On the other hand the stationary state σ_0 of LE is defined by the property

$$\mathcal{E}_\tau(\sigma_0) = \sigma_0 \quad \forall \tau \geq 0, \quad (89)$$

while the characteristic time τ_E by the property

$$\mathcal{E}_{\tau \geq \tau_E}(\Theta) = \sigma_0 \text{Tr} \{\Theta\}, \quad (90)$$

for all bounded operator Θ of \mathcal{H}_{LE} .

The resulting transformation in $\mathcal{H}_C \otimes \mathcal{H}_{LE}$ is a time ordered product of interactions U_j and relaxation processes \mathcal{E}_{τ_j} (see Fig. 6). Assuming LE to be initially in the stationary state σ_0 this gives a discrete family of CPTP maps $\{\Phi_s^{(n)}\}_n$ for an n carriers state $\rho^{(n)}$,

$$\rho^{(n)} \mapsto \Phi_s^{(n)}(\rho^{(n)}) = \text{Tr}_{LE} \{ \mathcal{U}_n \circ \mathcal{E}_{\tau_{n-1}} \circ \mathcal{U}_{n-1} \circ \dots \circ \mathcal{E}_{\tau_1} \circ \mathcal{U}_1 (\rho^{(n)} \otimes \sigma_0) \}, \quad (91)$$

where $\mathcal{U}_j(\dots)$ stands for the unitary mapping $U_j(\dots)U_j^\dagger$ on $\mathcal{H}_{C_j} \otimes \mathcal{H}_{LE}$, and “o” indicates the composition of super-operators.

Because of the time ordering of Eq. (91) the output state of a carrier might depend on the input state of the carriers which precedes it in s but it is always independent from the input state of the carriers which follows it in the sequence. As a matter of fact Eq. (91) exhibits the general structure of causal memory channel with time-translation invariance see Sec. III.B.1 and (Kretschmann and Werner, 2005).

Now, assume Alice is producing a sequence s with intervals τ_j greater than or equal to the characteristic relaxation time τ_E of the dissipation process. In this case, after each interaction, the local environment LE has enough time to relax into the stationary configuration σ_0 before a new carrier begins interacting with it. Under this hypothesis Eqs. (90) and (91) yield

$$\Phi_s^{(n)} = \mathcal{N}^{\otimes n} \quad (92)$$

where \mathcal{N} is the CPTP map which transforms the density matrices ρ of a single carrier into

$$\mathcal{N}(\rho) = \text{Tr}_{LE} \{ \mathcal{U}(\rho \otimes \sigma_0) \}. \quad (93)$$

Equation (92) describes a memoryless configuration where the noise acts on the q_j independently.

Alternatively, suppose that only a subset Q_M of the transmitted carriers is used to encode messages to Bob. The remaining carriers (subset Q_P) are instead employed for perturbing LE in such a way that the q_j on which the messages are encoded have a better chance to reach Bob without being corrupted. A simple implementation of a noise attenuation scheme is such that the Q_P carriers are composed by uniform strings of n states ρ_0 separated by equal time intervals τ . The information is instead encoded a single carrier and the whole structure repeats after a relaxation time τ_E . In this configuration the transformation of the Q_A carriers which comes from solving Eq. (91) can be computed as follows. First, one determines the modified state σ_n of LE which arises from the interactions with the B carriers. This is accomplished by solving the set of coupled equations

$$\begin{cases} \sigma'_j = \text{Tr}_C \{ \mathcal{U}(\rho_0 \otimes \sigma_j) \}, & j = 0, 1, \dots, n-1. \\ \sigma_{j+1} = \mathcal{E}_\tau(\sigma'_j), \end{cases} \quad (94)$$

The density matrix σ_n which results from (94) is then used to determine the output state of the A carriers according to the equation

$$\bar{\mathcal{N}}(\rho) = \text{Tr}_{LE} \{ \mathcal{U}(\rho \otimes \sigma_n) \}. \quad (95)$$

The transformation (95) is in general different from Eq. (93) and depends explicitly on the parameters n , τ and ρ_0 , so it can be regarded as a memory channel.

The correlated dephasing quantum channel can be considered as an example of the above described dynamical model. We assume the carrier-LE interaction U_j to be

to a control-unitary such that when the carrier is in $|0\rangle_{q_j}$ nothing happens to LE, while when q_j is in $|1\rangle_{q_j}$ the environment undergoes to the transformation

$$\begin{pmatrix} g & \sqrt{1-g^2} \\ \sqrt{1-g^2} & -g \end{pmatrix}, \quad (96)$$

with $g \in [0, 1]$. Moreover, one will assume the relaxation process $\{\mathcal{E}_\tau\}_\tau$ acting on LE to be described by amplitude damping maps which takes the state $|1\rangle_{LE}$ to $|0\rangle_{LE}$ with probability $1-\eta(\tau)$ where $\eta(\tau) \in [0, 1]$ is a non increasing function of τ with characteristic time τ_E , i.e.

$$\begin{aligned} \mathcal{E}_\tau(|0\rangle_{LE}\langle 0|) &= |0\rangle_{LE}\langle 0| \\ \mathcal{E}_\tau(|1\rangle_{LE}\langle 1|) &= \eta(\tau) |1\rangle_{LE}\langle 1| + (1-\eta(\tau)) |0\rangle_{LE}\langle 0| \\ \mathcal{E}_\tau(|0\rangle_{LE}\langle 1|) &= \sqrt{\eta(\tau)} |0\rangle_{LE}\langle 1|. \end{aligned} \quad (97)$$

In this example the stationary state σ_0 of LE is hence $|0\rangle_{LE}$. The parametrization of the memory effect is given by $\eta(\tau)$, with $\eta = 0$ corresponding to the memoryless case (fast environment relaxation, say $\tau \gg \tau_E$) and $\eta = 1$ corresponding to perfect memory case (no environment relaxation, say $\tau \ll \tau_E$). Under the above conditions, it is possible to show that both the map \mathcal{N} of the memoryless case and the map $\bar{\mathcal{N}}$ of the noise attenuation protocol correspond to a phase damping channel \mathcal{P}_g where the coherence terms of the input qubit ρ are degraded by the factor $g = 1 - 2p_z$ with p_z the probability of z error, i.e.

$$\begin{aligned} \mathcal{P}_g(|0\rangle_C\langle 0|) &= |0\rangle_C\langle 0|, & \mathcal{P}_g(|1\rangle_C\langle 1|) &= |1\rangle_C\langle 1|, \\ \mathcal{P}_g(|0\rangle_C\langle 1|) &= g |0\rangle_C\langle 1|. \end{aligned} \quad (98)$$

In particular Eq. (93) gives $\mathcal{N} = \mathcal{P}_g$. On the other hand, Eq. (95) gives $\bar{\mathcal{N}} = \mathcal{P}_{\bar{g}}$ where \bar{g} is a complicated expression of the parameters ρ_0 , n and τ . By appropriately selecting the values of the above quantities one can make $\bar{\mathcal{N}}$ less noisy than \mathcal{N} by having $\bar{g} > g$.

In the case of the phase damping channels the capacities can be explicitly computed. For instance, since the noise does not affect the populations associated with the computational basis, the classical capacity of the phase damping channel (98) $C(\mathcal{P}_g) = 1$.

On the other hand the quantum capacity of a phase damping channel (98) is equal to

$$Q(\mathcal{P}_g) = 1 - H(p_z) = 1 - H(1/2 + g/2), \quad (99)$$

where H is the binary entropy. We hence have, for $\bar{g} > g$, $Q(\mathcal{P}_{\bar{g}}) \geq Q(\mathcal{P}_g)$, i.e. enhanced quantum capacity by memory effects.

Markovian correlated noise can be derived from the mapping (91) by properly choosing the transformation \mathcal{E}_{τ_j} . Consider the case in which for sufficiently big τ the map \mathcal{E}_τ describes a decoherent process of LE where, given $\{|\ell\rangle_{LE}\}$ an orthonormal basis of \mathcal{H}_{LE} , one has

$$\mathcal{E}_\tau(|\ell\rangle_{LE}\langle \ell'|) = \delta_{\ell,\ell'} |\psi_\ell(\tau)\rangle_{LE}\langle \psi_\ell(\tau)|, \quad (100)$$

with the vectors $\{|\psi_\ell(\tau)\rangle_{LE}\}_\ell$ being not necessarily orthogonal. The condition (89) can then be satisfied by identifying σ_0 with one element of the selected basis (say $|\ell_0\rangle_{LE}$), and imposing $|\psi_\ell(\tau \geq \tau_E)\rangle_{LE} = |\ell_0\rangle_{LE}$ for all ℓ . In this case the mapping (91) can be expressed in terms of the operators

$$\tilde{A}_{\ell_1} = {}_{LE}\langle \ell_1 | U_1 | \ell_0 \rangle_{LE} \quad (101)$$

$$\tilde{A}_{\ell_{j+1}} = {}_{LE}\langle \ell_{j+1} | U_{j+1} | \psi_{\ell_j}(\tau_j) \rangle_{LE}, \quad (102)$$

which act, respectively, on the Hilbert space \mathcal{H}_{C_1} and $\mathcal{H}_{C_{j+1}}$ for $j = 1, \dots, n-1$. They allow us to define the probability distribution

$$p_{\ell_1} := \text{Tr}_{q_1} \left\{ \tilde{A}_{\ell_1}^\dagger \tilde{A}_{\ell_1} \right\} \quad (103)$$

and the conditional probabilities

$$p_{\ell_{j+1}|\ell_j} := \text{Tr}_{q_{j+1}} \left\{ \tilde{A}_{\ell_{j+1}}^\dagger \tilde{A}_{\ell_{j+1}} \right\}. \quad (104)$$

Using these quantities Eq. (91) can be finally expressed in compact Markovian form,

$$\begin{aligned} \Phi_s^{(n)}(\rho^{(n)}) &= \sum_{\ell_1, \dots, \ell_n} p_{\ell_1} p_{\ell_2|\ell_1} \cdots p_{\ell_n|\ell_{n-1}} \\ &\times A_{\ell_n} \cdots A_{\ell_2} A_{\ell_1} \rho^{(n)} A_{\ell_1}^\dagger A_{\ell_2}^\dagger \cdots A_{\ell_n}^\dagger, \end{aligned} \quad (105)$$

with $A_{\ell_1} = \tilde{A}_{\ell_1}/\sqrt{p_{\ell_1}}$ and $A_{\ell_{j+1}} = \tilde{A}_{\ell_{j+1}}/\sqrt{p_{\ell_{j+1}|\ell_j}}$.

D'Arrigo, Benenti and Falci, 2007 have considered dephasing channel along this line, i.e. with Markovian correlations. Specifically, they have taken the Kraus operators as $A_{\ell_j=0} = \mathbb{I}, A_{\ell_j=1} = \sigma_z, \forall j$. Moreover, the conditional probability have been assumed as

$$p_{\ell_j|\ell_{j-1}} = (1 - \mu) p_{\ell_j} + \mu \delta_{\ell_j, \ell_{j-1}}, \quad (106)$$

where $\mu \in [0, 1]$ measures the partial memory of the channel.

Clearly the classical capacity is $C = 1$ also in this case.

In order to compute the quantum capacity, one evaluates the maximum of coherent information $J(\Phi^{(n)}, \rho^{(n)})$ by choosing $\rho^{(n)} = I_n/2^n$. In this case $S(\Phi^{(n)}(\rho^{(n)})) = S(\rho^{(n)}) = n$. Furthermore the entropy exchange equals $S(W)$, where the density operator W has components $W_{i_1 \dots i_n, i'_1 \dots i'_n} = \text{Tr}(A_{i_1 \dots i_n} \rho^{(n)} A_{i'_1 \dots i'_n}^\dagger)$. Here W is diagonal and

$$S(W) = - \sum_{\{i_k\}} p_{\{i_k\}} \log_2 p_{\{i_k\}} \equiv H(X_1, \dots, X_n), \quad (107)$$

where $H(X_1, \dots, X_n)$ is the Shannon entropy of the collection of random variables X_1, \dots, X_n (characterized by the joint probabilities $p_{i_1 \dots i_n}$). For a stationary Markov chain, one has (Cover and Thomas, 1991)

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} H(X_1, \dots, X_N) &= H(X_2|X_1) \\ &= p_0 H(q_0) + p_z H(q_z), \end{aligned} \quad (108)$$

where $q_{0,z} = (1 - \mu)p_{0,z} + \mu$ are the conditional probabilities that the channel acts on two subsequent qubits via the same Pauli operator, and $H(q_0), H(q_z)$ are binary Shannon entropies. Therefore, the quantum capacity is finally given by

$$Q = 1 - p_0 H(q_0) - p_z H(q_z). \quad (109)$$

Interestingly enough, Q increases when considering an higher degree of memory. In particular, the memoryless dephasing channel capacity, $Q = Q_1 = 1 - H(p_0)$, is recovered for $\mu = 0$, while for $\mu = 1$ (perfect memory) the channel is asymptotically noiseless, i.e. $Q = 1$ (Bowen and Mancini, 2004).

A microscopic model for correlated dephasing has been introduced in D'Arrigo, Benenti and Falci, 2007 in terms of a spin-boson model, where quantum information is encoded in a train of qubits and a single bosonic mode represents the memory system. Lower bounds for the quantum capacity of a qubit memory channel with both correlated dephasing and damping have been evaluated numerically starting from a microscopic spin-boson model with Jaynes-Cummings interaction in the presence of strong dephasing noise (Benenti, D'Arrigo and Falci, 2009).

Plenio and Virmani, 2007, 2008 have considered another model of dephasing memory channel for qubits. It can be traced back to the scenario introduced in (Giovannetti and Mancini, 2005) and schematized in Fig. 3 where each individual information carrier (a qubit in this case) interacts with a correspondent environment particle, the correlations being established by the environment multi-particle state. Specifically they have considered the case where the two-particle (two-qubit) interaction is defined by a controlled-phase gate, the environmental particle being the controller qubit that determines which unitary transformation will be applied to the carrier. As a consequence the joint state of the carriers gets transformed through mixtures of random sequences of identity and σ_z operators, each sequence being characterized by a (correlated) probability which depends upon the diagonal elements of the environment initial state.

The interesting feature of this model is that it allows to write down explicit formulae for the associated capacities for the channel in terms of properties of the many-body environment that share a close relationship with thermodynamical quantities. In particular the quantum capacity can be expressed in terms of the regularized diagonal entropy of the system environment, i.e.

$$Q = 1 - \lim_{n \rightarrow \infty} \frac{S(\text{diag}(\rho_{env}))}{n}, \quad (110)$$

where $\text{diag}(\rho_{env})$ is the environmental state in the computational basis after eliminating all off-diagonal elements [note that the coding argument used in order to arrive to Eq. (110) was been also independently shown by (Hamada, 2002)]. For the special case in which the initial state of the environment is described by a classically

correlated many-body system, the last term on the rhs of Eq. (110) is exactly the entropy. Hence, the capacity is given by

$$Q = 1 - \left(1 - \beta \frac{\partial}{\partial \beta}\right) \lim_{n \rightarrow \infty} \frac{1}{n} \log_2 Z_n, \quad (111)$$

where Z_n is the partition function for n environment spins, and β is the associated inverse temperature. In other words, one can exploit results from classical statistical physics in order to compute the capacity, as shown by Eq. (111). If on the other hand, the state ρ_{env} involve also quantum correlations, Eq. (110) reduces to the entropy obtained when every environment qubit is completely dephased. Before computing it, one can notice that (Wolf *et al.*, 2006) showed the existence of Hamiltonians exhibiting quantum phase transitions and with ground states being Matrix Product States (MPS) involving only rank-1 matrices. Hence, it can be shown that the diagonal elements of such MPSs are equal to the probability \wp of microstates in corresponding classical Ising chains.

For a sake of simplicity, one does focus on a translationally invariant MPS for a 1D system of 2-level particles, in the case of periodic boundary conditions. This environmental state is characterized by two matrices Q_0 and Q_1 and is given by the following expression $|\psi\rangle = \sum_{i_1, \dots, i_n} \text{tr}\{Q_{i_1} \dots Q_{i_n}\} |i_1 \dots i_n\rangle$. Then, by dephasing each qubit, the resulting unnormalized state is

$$\rho = \sum_{i_1, \dots, i_n} \text{tr} \left[\prod_{k=1}^n (Q_{i_k} \otimes \bar{Q}_{i_k}) \right] |i_1 \dots i_n\rangle \langle i_1 \dots i_n|. \quad (112)$$

It is possible to show that, if $|i_1 \dots i_n\rangle$ has l occurrences of 0 and $n - l$ of 1, and k boundaries between 0s and 1s blocks, then the corresponding diagonal elements of ρ are

$$\wp(l, n - l, k) = a^l b^{n-l} c^k / \mathfrak{N}(n), \quad (113)$$

with a (resp. b) being the eigenvalue of $Q_0 \otimes \bar{Q}_0$ (resp. $Q_1 \otimes \bar{Q}_1$), c being the eigenvalue of $(Q_0 \otimes \bar{Q}_0)(Q_1 \otimes \bar{Q}_1)/(ab)$ and $\mathfrak{N}(n)$ a normalization factor.

Therefore, by exploiting this connection, one can easily compute the limit in Eq.(110) by using well known many-body physics methods. Fig. 7 shows the case of the following Hamiltonian

$$\sum_i 2(g^2 - 1)\sigma_z^i \sigma_z^{i+1} - (1 + g)^2 \sigma_x^i + (g - 1)^2 \sigma_z^i \sigma_x^i \sigma_z^{i+1}. \quad (114)$$

In this case, one knows that the ground state is a rank-1 MPS which possesses a non-standard ‘phase transition’ at $g = 0$, where indeed some correlation functions are non-differentiable (though continuous) and the ground state energy is analytic (Wolf *et al.*, 2006).

B. Continuous memory channels

Bosonic or continuous-variable (CV) systems are often used to encode classical and quantum information,

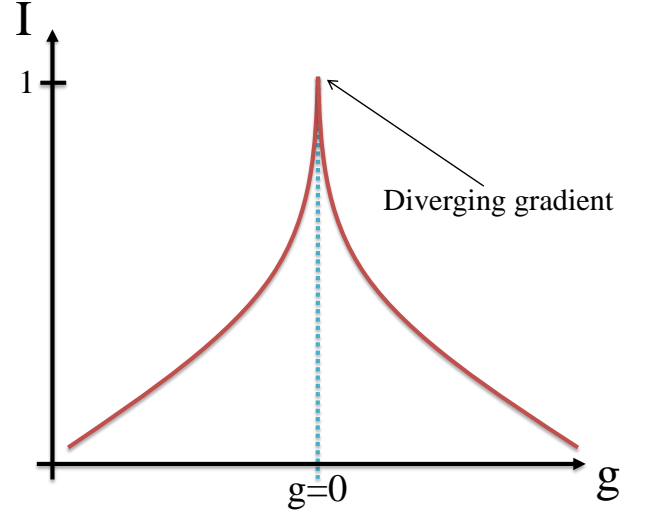


FIG. 7 Sketch of the capacity behaviour in the case of an environment given by the ground state of the Hamiltonian (114). Notice the divergent gradient near the ‘phase transition’, i.e. at $g = 0$ (Plenio and Virmani, 2007).

e.g., photons transmitted in optical fibers, wave guides, free-space, etc. (Braunstein and van Loock, 2005; Eisert and Plenio, 2003; Ferraro, Olivares and Paris, 2005; Weedbrook *et al.*, 2012). Hence, an important class of quantum channels are those acting on CV quantum systems (Caves and Drummond, 1994). For describing CV systems, one introduces a discrete set of bosonic oscillators, typically a set of normal modes of the electromagnetic field, described by the canonical pairs $\{x_k, y_k\}_{k=1, \dots, n}$, the generalized “position” and “momentum”, from which one defines the vector of phase-space variables $\mathbf{z} := (x_1, y_1, \dots, x_n, y_n)^t$ and the ladder operators $\{a_k, a_k^\dagger\}$ obeying the canonical commutation relations $[a_h, a_k^\dagger] = \delta_{hk}$. The state of the bosonic system is described by a density operator ρ , or equivalently by the associated characteristic function $\chi(\mathbf{z}) = \text{Tr}[\rho V(\mathbf{z})]$, where $V(\mathbf{z}) = \exp\left[\frac{1}{\sqrt{2}} \sum_k (x_k + \iota y_k) a_k^\dagger - (x_k - \iota y_k) a_k\right]$ denotes the n -mode Weyl operator (ι is the imaginary unit).

A remarkable class of states is that of the so-called Gaussian states (Eisert and Plenio, 2003; Ferraro, Olivares and Paris, 2005). These are the states having a Gaussian characteristic function, $\chi(z) = \exp(\iota \mathbf{m}^t \mathbf{z} - \frac{1}{2} \mathbf{z}^t \gamma \mathbf{z})$, where \mathbf{m} is the vector of first moments and γ is the covariance matrix (CM) of the phase-space variables. The canonical commutation relations lead to the uncertainty relations, which in terms of the CM are expressed by the matrix inequality $\gamma - \iota \Omega / 2 \geq 0$, where

$$\Omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (115)$$

is the matrix representation of the phase-space canonical symplectic form.

Going back to quantum channels in CV systems, up to now attention has been mainly devoted to the study of Gaussian channels. Gaussian channels have the characterizing properties of transforming Gaussian states into Gaussian states. The action of a n -mode Gaussian channels, in the Schrödinger picture, transforms a state with characteristic function $\chi(\mathbf{z})$ into

$$\chi'(\mathbf{z}) = \chi(X^{(n)}\mathbf{z})f(\mathbf{z}), \quad (116)$$

where $X^{(n)}$ is a matrix inducing a linear transformation on the $2n$ -dimensional phase-space vector \mathbf{z} , and the function $f(\mathbf{z})$ is Gaussian, i.e. $f(\mathbf{z}) = \exp(i\mathbf{d}^{(n)t}\mathbf{z} - \frac{1}{2}\mathbf{z}^t Y^{(n)}\mathbf{z})$ (Holevo and Werner, 2001). The linear term proportional to the vector $\mathbf{d}^{(n)}$ accounts for a translation (displacement) of the mean \mathbf{m} , while the quadratic term proportional to the matrix $Y^{(n)}$ adds a term to the CM. Gaussian transformations which are also unitary are characterized by the property that $X^{(n)}$ is a symplectic matrix (i.e., preserving the canonical form, $X^{(n)t}\Omega X^{(n)} = \Omega$), and $Y^{(n)} = 0$. A n -mode Gaussian channel is hence characterized by the triad $(\mathbf{d}^{(n)}, X^{(n)}, Y^{(n)})$. The composition of two Gaussian channels with associated triads $(\mathbf{d}_1^{(n)}, X_1^{(n)}, Y_1^{(n)})$ and $(\mathbf{d}_2^{(n)}, X_2^{(n)}, Y_2^{(n)})$ yields $(X_2^{(n)}\mathbf{d}_1^{(n)} + \mathbf{d}_2^{(n)}, X_2^{(n)}X_1^{(n)}, X_2^{(n)}Y_1^{(n)}X_2^{(n)t} + Y_2^{(n)})$. It follows that, by applying suitable Gaussian unitaries at the input and output of the channel, one can always reduce the channel in a canonical form, in which $\mathbf{d}^{(n)} = 0$, and the matrices $X^{(n)}, Y^{(n)}$ take a particular symmetric form. For the case of channels acting on one or two modes, the reduction to canonical forms allows the classification of Gaussian quantum channels according to invariance under unitary transformations (Caruso and Giovannetti, 2006; Caruso *et al.*, 2008; Holevo, 2007). If one takes the one-mode channel as a reference point, representing a single use of the channel, a memory channel is characterized by a sequence of triads $(\mathbf{d}^{(n)}, X^{(n)}, Y^{(n)})$ such that either $\mathbf{d}^{(n)} \neq \bigoplus_{k=1}^n \mathbf{d}^{(1)}$ or $X^{(n)} \neq \bigoplus_{k=1}^n X^{(1)}$, $Y^{(n)} \neq \bigoplus_{k=1}^n Y^{(1)}$.

Due to the fact that the carrier Hilbert space is infinite-dimensional it turns out that the capacity of CV channel can be infinite. It is hence meaningful to introduce a notion of constrained capacity. Most natural choices are to impose a constraint on the mean energy, or on the mean number of bosonic excitations. Given n uses of the channel, the constraint on the mean excitation number reads

$$\frac{\text{Tr}(\gamma) + |\mathbf{m}|^2}{2n} \leq N + \frac{1}{2}, \quad (117)$$

where N is the maximum mean excitation number per input mode.

For the case of Gaussian channels, taking also into account that the energy and the excitation number are quadratic in the canonical operators, it is natural to conjecture that the optimization is saturated by Gaussian input states. However, the optimality of Gaussian

input states has been rigorously proven only for some cases (Giovannetti, *et al.*, 2004; Wolf, Pérez-García and Giedke, 2007).

If one circumvents this problem and restricts the optimization to Gaussian states and/or to states which are separable among different channel uses, the obtained quantities are lower bounds on the capacity. The result of the optimization problem under the restriction of Gaussian input state is also called the *Gaussian capacity*. These lower bounds, together with upper bounds, were first considered in Holevo and Werner, 2001. Restricting the optimization to Gaussian input states provides dramatic simplification in the calculation. In fact, a simple formula exists for the von Neumann entropy of Gaussian states (Holevo, *et al.*, 1999). Given a n -mode Gaussian state with characteristic function $\chi(z) = \exp(i\mathbf{m}^t\mathbf{z} - \frac{1}{2}\mathbf{z}^t\gamma\mathbf{z})$, its von Neumann entropy is a function of the CM only:

$$S = \sum_{k=1}^n g(\nu_k - 1/2), \quad (118)$$

where

$$g(x) := \begin{cases} (x+1)\log(x+1) - x\log x & \text{for } x > 0, \\ 0 & \text{for } x \leq 0, \end{cases} \quad (119)$$

and $\{\nu_k\}_{k=1,\dots,n}$ are the symplectic eigenvalues of the CM. For a n mode Gaussian state with CM γ , the symplectic eigenvalues are computed from the eigenvalues of $\gamma \bigoplus_{k=1}^n \Omega$, where Ω is the symplectic form, Eq. (115). Explicitly, the eigenvalues of $\gamma \bigoplus_{k=1}^n \Omega$ are $\{\pm i\nu_k\}_{k=1,\dots,n}$. Since Gaussian channels maps Gaussian input states into Gaussian output states, all the entropic functions involved in the calculation of capacities, as the Holevo function, the coherent information, and the quantum mutual information, can be computed starting from this expression for the von Neumann entropy.

Among Gaussian memory channels, one can identify a subclass of channels for which the memory effects can be *unraveled*. That is, by applying suitable unitary encoding and decoding transformations, n uses of such channels are unitary equivalent to the n single-mode channels used in parallel. By applying known results for the memoryless setting one may then compute the capacities of the memory channel or estimate lower bounds in terms of the Gaussian capacities.

Such a unitary mapping from n uses of a Gaussian memory channel to n parallel uses of independent single-mode channels was first considered in (Cerf *et al.*, 2005, 2006; Giovannetti and Mancini, 2005), and then applied for estimating the communication capacities of Gaussian memory channels in several settings (Lupo, Memarzadeh and Mancini, 2009; Lupo, Pilyavets and Mancini, 2009; Lupo, Giovannetti and Mancini, 2010,b; Schäfer, Karpov and Cerf, 2009). A formal definition of the class of memory channels that can be unraveled first appeared in (Lupo and Mancini, 2010).

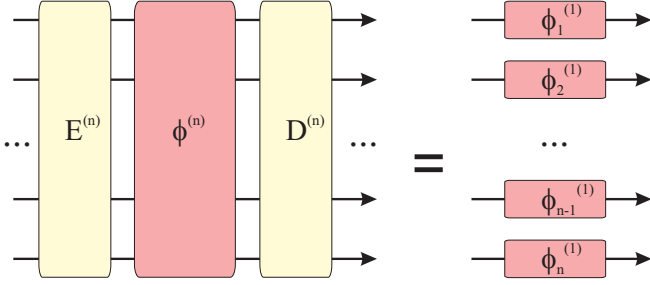


FIG. 8 Unraveling of n uses of a memory channel. Each horizontal line indicates one bosonic mode, propagating from the left to the right. $\phi^{(n)}$ denotes n uses of the memory channel. $E^{(n)}$ and $D^{(n)}$ are pre-processing and post-processing Gaussian unitaries. $\phi_k^{(1)}$'s are one-mode Gaussian channels.

For a given n , let us consider n uses of a Gaussian memory channel, denoted $\phi^{(n)}$, associated with the triad $(\mathbf{d}^{(n)}, X^{(n)}, Y^{(n)})$. A memory channel can be unraveled if there exist unitary transformations $E^{(n)}$, $D^{(n)}$, acting on n modes, such that $D^{(n)}\phi^{(n)}E^{(n)} = \bigotimes_{k=1}^n \phi_k^{(1)}$, that is, n uses of the memory channel are unitary equivalent to the tensor product of n independent, but not necessarily identical, single-mode Gaussian channel. Recalling that a n -mode Gaussian unitary transformation is identified by a triad $(\mathbf{d}^{(n)}, S^{(n)}, 0)$ where $\mathbf{d}^{(n)}$ is a displacement vector and $S^{(n)}$ is a symplectic matrix, it follows that the memory channel can be unraveled if and only if there exist, for any n , symplectic matrices $S_E^{(n)}$, $S_D^{(n)}$ such that $S_D^{(n)} X^{(n)} S_E^{(n)} = \bigotimes_{k=1}^n X_k^{(1)}$, and $S_D^{(n)} Y^{(n)} [S_D^{(n)}]^\dagger = \bigotimes_{k=1}^n Y_k^{(1)}$. Since the application of unitary transformations cannot change the capacities of the channel, they can be equivalently computed for the unraveled channel, in which each input mode is transformed independently (although in general not-identically). This mapping is depicted in Fig. 8. If one is interested in the calculation of constrained capacities, then one has to take in account how the input energy is changed by the encoding unitary $E^{(n)}$. The most relevant case is when the input photon-number remains unchanged, that is, $\sum_{k=1}^n a_k^\dagger a_k = E^{(n)\dagger} \sum_{k=1}^n a_k^\dagger a_k E^{(n)}$. This is the case when the encoding unitary is a so-called linear passive transformation. For optical realization, such transformations are implemented by a network of beam-splitters and phase-shifters (see e.g. (Ferraro, Olivares and Paris, 2005)).

Let us now review results on solvable models belonging to the two most relevant families of models of CV memory channels.

Additive-noise channel, or channel with thermal-like noise. In the Schrödinger picture, the input state over n channel uses is transformed to

$$\phi^{(n)}(\rho) = \int d\xi^n P_{\xi_1, \dots, \xi_n} \left[\bigotimes_k V(\xi_k) \right] \rho \left[\bigotimes_k V^\dagger(\xi_k) \right], \quad (120)$$

where $V(\xi_k)$ is the Weyl operator for the k -th input

mode. If P is factorized as function of ξ_1, \dots, ξ_n , the channel is memoryless, otherwise, it is a channel with memory. This class of Gaussian channels is characterized by having $X^{(n)} = 1$. Without loss of generality, one put $\mathbf{d}^{(n)} = 0$. For the case of memoryless thermal noise, the matrix $Y^{(n)}$ is scalar, $Y^{(1)} = N_{\text{th}}$, with $N_{\text{th}} \geq 0$, and the Gaussian quantum and classical capacities can be computed exactly (Holevo and Werner, 2001). The memory channel considered in (Cerf *et al.*, 2005, 2006) belongs to this class. The latter is defined for two channel uses, represented by two bosonic modes, which are affected by correlated additive noise. A generalization of this model to the case of more than two channel uses was introduced in (Schäfer, Karpov and Cerf, 2009), based on a multi-mode Markovian extension of the two-mode correlated noise. The additive noise has a Markovian structure with associated conditional probability $P_{\xi_{k+1}|\xi_k}$. By a suitable choice of the conditional probability and of the distribution of the initial noise variable ξ_0 , one gets a zero-mean Gaussian process describing the additive noise, with an associated triad of the form $(0, 1, Y^{(n)})$. Here, one considers the models introduced in (Schäfer, Karpov and Cerf, 2009) and (Lupo, Memarzadeh and Mancini, 2009), for which $Y^{(n)}$ has the form of a block-matrix

$$Y^{(n)} = \begin{pmatrix} M_{11} & M_{12} & \dots & M_{1n} \\ M_{21} & M_{22} & \dots & M_{2n} \\ \vdots & \vdots & \dots & \vdots \\ M_{n1} & M_{n2} & \dots & M_{nn} \end{pmatrix}, \quad (121)$$

with $M_{ij} = \sigma \phi^{|i-j|} Z^\ell$, and $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The parameters $\sigma \geq 0$ and $\phi \in [0, 1)$ determine the noise strength and correlations and the index $\ell = 0, 1$ determines whether the noise is symmetric or anti-symmetric under exchange of the phase-space coordinates. For $\ell = 0$ one gets the model considered in (Lupo, Memarzadeh and Mancini, 2009), and for $\ell = 1$ that considered in (Schäfer, Karpov and Cerf, 2009). The memory channel is unraveled by the application of unitary pre-processing and post-processing, which are realized as multiport beam-splitter. For the case $\ell = 0$, the channel can be unraveled for any n . After unraveling, the channel appears as the tensor product of single-mode channel with thermal noise, for which the optimal input states are coherent states, which are separable among channel uses. On the other hand, for $\ell = 1$, the memory channel cannot be unraveled for finite values of n . However, it can be unraveled in the limit $n \rightarrow \infty$, hence allowing the computation of the channel capacities. The unraveled channel is the tensor product of single-mode channel with additive noise which is asymmetric in the phase-space. Due to this asymmetry, the optimal input states for the classical capacity are multimode squeezed states, which are entangled among different channel uses. Differently from the case of discrete-variable memory channels (Macchiavello, Palma and Virmani, 2004), there is no transition in these Gaussian models: the optimal input states are either separable

or entangled according to the model symmetries (Cerf *et al.*, 2005, 2006; Lupo and Mancini, 2010). As entangled states cannot be prepared locally, it is crucial to identify suboptimal input states that can be prepared efficiently. This issue was considered in (Schäfer, Karpov and Cerf, 2011), where it was shown that encoding classical information via Gaussian matrix-product states (Adesso and Ericsson, 2006; Schuch, Cirac and Wolf, 2008), which can be efficiently prepared, may allow to achieve a reliable communication rate close to the channel capacity. Finally, an analysis of correlated additive Gaussian channels beyond the case of Markovian correlations was presented in (Schäfer, Karpov and Cerf, 2011).

Another family of memory channels has been first introduced in (Giovannetti and Mancini, 2005), belonging to the class of *lossy bosonic channels*. Upon n uses of the memory channel, n input modes interact mode-wise with n environmental modes by a beam-splitter transformation with given transmissivity and phase. This is a realization of the general scheme depicted in Fig. 3, where each horizontal line represents an ingoing bosonic mode, each vertical line an environmental mode, and the boxes are the beam-splitters. Clearly, no correlations exist if the environmental modes are in a product state. In particular, a memoryless Gaussian channel is obtained for a product of identical Gaussian states (Pilyavets, Lupo and Mancini, 2009), e.g. for the case of the vacuum state the memoryless purely lossy bosonic channel is recovered (Giovannetti, *et al.*, 2004; Wolf, Pérez-García and Giedke, 2007), and for a thermal state one recovers the memoryless lossy bosonic channel with thermal noise (Giovannetti, *et al.*, 2010). On the other hand, memory effects arise whenever there are correlations among the environmental modes. The environmental state considered in (Giovannetti and Mancini, 2005) is a n -mode squeezed vacuum, or a squeezed thermal state, in which multimode squeezing introduces correlations among channel uses, described by a density operator of the form:

$$\rho_{\text{env}} = \mathcal{S} \rho_{\text{th}} \mathcal{S}^\dagger, \quad (122)$$

where ρ_{th} is a n -mode thermal state, with mean photon number equal to N_{th} , and

$$\mathcal{S} = \exp \left(\sum_{k,k'} \bar{\beta}_{kk'} v_k v_{k'} - \beta_{kk'} v_k^\dagger v_{k'}^\dagger \right), \quad (123)$$

with $\{v_k, v_k^\dagger\}$ being the canonical ladder operators associated to the k -th environment mode. n uses of the channel are associated to a triad $(0, \sqrt{\eta}, (1-\eta)\gamma_{\text{env}})$, where γ_{env} is the CM of the n -mode environment state. The CM of the state in Eq. (122) is of the form

$$\gamma_{\text{env}} = \left(N_{\text{th}} + \frac{1}{2} \right) \mathcal{S} \mathcal{S}^\dagger, \quad (124)$$

where \mathcal{S} is the symplectic matrix being the phase-space representation of the unitary in (123). As first noticed

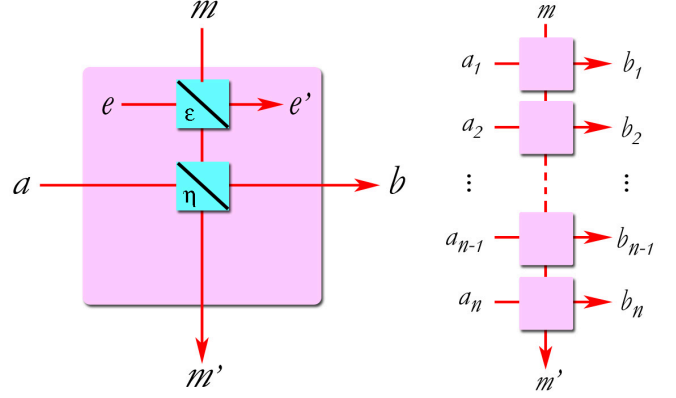


FIG. 9 (Color online) Left: a single use of the lossy bosonic memory channel (Lupo, Giovannetti and Mancini, 2010). Right: the n -fold concatenation of the memory channel: photons entering in the k -th input mode a_k can only emerge in the output ports $b_{k'}$ with $k' \geq k$.

in (Giovannetti and Mancini, 2005) this channel model is unitary equivalent to the tensor product of independent, but not-identical, one-mode lossy bosonic Gaussian channel, that is, it can be unraveled. Giovannetti and Mancini, 2005 provided upper and lower bounds on the classical capacity of the memory channel in terms of the capacity of the memoryless single-mode lossy bosonic channel after replacing the input mean photon-number N with an effective values $\bar{N} > N$ (for obtaining the upper bound) and $N' < N$ (for obtaining the lower bound). It has been shown in (Lupo, Pilyavets and Mancini, 2009), using a class of environmental states (e.g. that introduced in (Pilyavets, Zborovskii and Mancini, 2008)), that there exists a range of parameters for which the memory channel can be unraveled with the use of unitary pre-processing transformation which do preserve the input energy. If the environment is in a pure state, a closed formula has been obtained for the constrained classical capacity:

$$C_N = g[\eta N + (1-\eta)\mathcal{M}], \quad (125)$$

where \mathcal{M} is a known function of the parameters of the environmental state. The same approach has been used in (Lupo, Pilyavets and Mancini, 2009) to evaluate bounds on the quantum and entanglement-assisted classical capacity. A different kind of model of *lossy bosonic channel* with memory was introduced in (Lupo, Giovannetti and Mancini, 2010). In this model the correlations among channel uses are not induced by the environmental correlations, but arise directly from the structure of the channel. As depicted in Fig. 9, the action of the channel upon n uses is defined by the concatenation of n identical unitary transformation coupling the input systems a_1, a_2, \dots, a_n with a collection of local environments e_1, e_2, \dots, e_n and the memory system m . Both the local environments and the memory system are represented by bosonic modes. Specifically the evolution of the k -th

input mode is obtained by a concatenation of two beam-splitter transformations, the first with transmissivity ϵ and the second with transmissivity η , as shown in Fig. 9. This results in a non-anticipatory intersymbol interference channel (see Sec. III.B.1 and III.B.3), in which previous input states affect the action of the channel on the current input (Bowen and Mancini, 2004). By varying the transmissivity parameters, the model may be reduced to a memoryless lossy bosonic channel (Giovannetti, *et al.*, 2004) (the input a_k only influences the output b_k), or to a channel with perfect memory (Bowen and Mancini, 2004) (all a_k interacts *only* with the memory mode m_1). For specific values of the parameters (that is, $\eta = 0$, $\epsilon = 1$) $\Phi^{(n)}$ describes a *quantum shift* channel (Bowen and Mancini, 2004), where each input state is replaced by the previous one. Such a model was extended in (Lupo, Giovannetti and Mancini, 2010b) to encompass memory effects in linear amplification processes.

In general, the calculation of the (Gaussian) capacities of these models of Gaussian memory channels can be divided in four steps: first the channel is mapped into the direct product of single-mode Gaussian channels, then the optimization of the relevant entropic function is performed mode-wise under constrained mean excitation number, then the distribution of the mean excitation number over the input modes is optimized, finally the asymptotic limit is considered. The optimization of the distribution of the mean excitation number leads to a *quantum water filling* solution for the capacity of the memory channel, where the way the mean excitation number is distributed over input modes is analogous to the way water distributes into a vessel (Cover and Thomas, 1991). Algorithms for the optimization were presented in (Pilyavets, Lupo and Mancini, 2009; Schäfer, Karpov and Cerf, 2011). The study of models of Gaussian memory channels that can be unraveled has also stimulated and motivated a deep analysis of the communication capacities of the single-mode (memoryless) Gaussian channel (Lupo, *et al.*, 2011; Pilyavets, Lupo and Mancini, 2009; Schäfer, Karpov and Cerf, 2010). In particular (Pilyavets, Lupo and Mancini, 2009) and (Schäfer, Karpov and Cerf, 2010) provided a characterization of one-mode Gaussian channels, respectively for the case of lossy channels and additive noise, in terms of the solutions of the optimization problem of computing the Gaussian classical capacity as a function of all the parameters entering the model.

VIII. QUANTUM CHANNELS FROM DYNAMICS WITH MEMORY

The formalism of quantum channels describes states changes according to the physical laws of quantum mechanics. Often, this states changes in open quantum systems are described continuously in time through the formalism of master equations. Within this approach, one usually resorts to the Born and Markov approxima-

tions (Breuer and Petruccione, 2002), respectively based on the assumption of weak system-environment coupling, and that the correlation times of the reservoir are much smaller than the typical time scale of the system dynamics. Hence, a different terminology is commonly used in this context, where the term Markovian becomes synonymous of memoryless. These approximations lead to the master equation

$$\frac{d}{dt}\rho(t) = \mathcal{L}\rho(t). \quad (126)$$

with the time-independent Liouvillian given by the Gorini-Kossakowski-Sudarshan-Lindblad expression (Gorini, Kossakowski and Sudarshan, 1976; Lindblad, 1976)

$$\mathcal{L}\rho(t) = -i[H, \rho(t)] + \sum_{\alpha} \left(V_{\alpha}\rho(t)V_{\alpha}^{\dagger} - \frac{1}{2}\{V_{\alpha}^{\dagger}V_{\alpha}, \rho\} \right), \quad (127)$$

where the operators H and V_{α} respectively describe the Hamiltonian and non-Hamiltonian dynamical terms. The solution of (126) can be written, for given initial condition $\rho(t_0)$, as $\rho(t) = \Lambda(t - t_0)\rho(t_0)$, where $\Lambda(t - t_0)$ defines a one-parameter family of dynamical maps which are completely positive and trace preserving, i.e. *quantum channels*. The dynamical map is formally given by $\Lambda(t - t_0) = e^{(t-t_0)\mathcal{L}}$. It obeys the homogeneous composition law

$$\Lambda(t_1)\Lambda(t_2) = \Lambda(t_1 + t_2), \quad (128)$$

for $t_1, t_2 \geq 0$.

A class of Markovian master equations of this form can be obtained as the continuous-time limit of a concatenation of identical system-bath interactions. These models, known as *collision models* (Alicki and Lendi, 1987; Scarani *et al.*, 2002; Terhal and DiVincenzo, 2000; Ziman and Buzek, 2005), are defined by the iterated unitary interactions of the system Q with n identical reservoirs $E = (e_1, \dots, e_n)$. This *cascade* process, depicted in Fig. 10, defines a quantum channel of the form

$$\Phi^n(\rho_Q) = \text{Tr}_E \left[U_{Qe_1} \cdots U_{Qe_n} (\rho_Q \otimes \Omega_E^{\otimes n}) U_{Qe_1}^{\dagger} \cdots U_{Qe_n}^{\dagger} \right], \quad (129)$$

where U_{Qe_j} 's indicate the identical unitary transformations coupling the system Q with the environmental systems. It is worth noticing that this expression shares the same structure of Eq. (17) after replacing M with Q and Q with E . This analogy is also apparent by comparing Fig. 10 with Fig. 1a. The unitary interaction can be explicitly written as $U_{Qe_j} = \exp(i g H_{Qe} \Delta t)$, where H_{Qe} is the system-environment Hamiltonian, g is the interaction strength and Δt is the interaction time. The continuous-time limit is hence defined by taking the limits $\Delta t \rightarrow 0$, and $n \rightarrow \infty$ in such a way that $n\Delta t \rightarrow T < \infty$, and

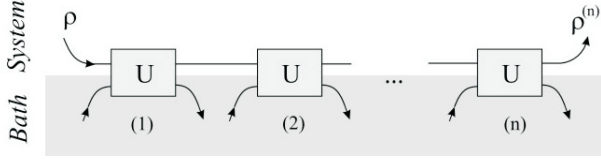


FIG. 10 The cascade structure of a collision model, defined by the concatenation of identical unitaries. Figure taken from (Scarani *et al.*, 2002)

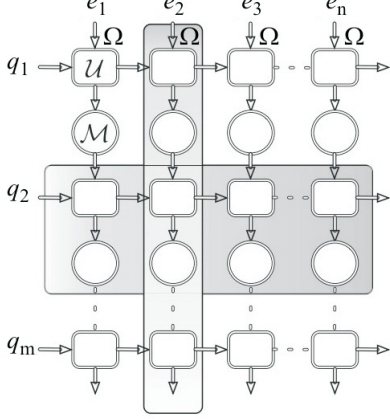


FIG. 11 The cascade structure leading to the master equation for correlated quantum channels discussed in (Giovannetti and Palma, 2012), described by Eq. (129). Each row corresponds to a single collision model (see Fig. 10), and each column correspond to a memory channel (see Fig. 1a)

$g^2 \Delta t \rightarrow \gamma < \infty$. In this limit the system evolution between time 0 and T is described by a Markovian master equation of the form (127). A more general framework has been recently introduced by (Giovannetti and Palma, 2012), where a collection of m quantum systems, $Q = (q_1, \dots, q_m)$, is considered. The system q_j repeatedly interacts with a corresponding collection of n environmental systems, $E_j = (e_1^j, \dots, e_n^j)$, according to Eq. (129). Furthermore, memory effects are included in this setup by cascading the environmental systems as depicted in Fig. 11, where an internal dynamics of the environmental systems is modeled by introducing an additional CPTP map \mathcal{M} . In the continuous-time limit, the overall dynamics of the multipartite system Q is still described by a Markovian master equation, while non Markovian effects appear in the reduced dynamics of the individual systems q_j .

A. Non Markovian master equations

The first and simplest generalization of the dynamical equation (126) is obtained by introducing a time-dependent Liouvillian $\mathcal{L}(t)$ admitting the representation (127), but with time-dependent operators, $H(t)$ and $V_\alpha(t)$. Hence, the time-dependent equation for the dy-

namical map $\Lambda(t, t_0)$

$$\frac{d}{dt} \Lambda(t, t_0) = \mathcal{L}(t) \Lambda(t, t_0), \quad \Lambda(t_0, t_0) = \text{id}, \quad (130)$$

has formal solution

$$\Lambda(t, t_0) = \text{T exp} \left(\int_{t_0}^t \mathcal{L}(\tau) d\tau \right), \quad (131)$$

where T denotes time-ordering. Differently from the time-homogeneous case, the explicit dependence on time implies that the dynamical map $\Lambda(t, t_0)$ is no more a function of ' $t - t_0$ ' only. Notwithstanding, it still satisfies the inhomogeneous composition law

$$\Lambda(t, s) \cdot \Lambda(s, t_0) = \Lambda(t, t_0), \quad (132)$$

for any $t \geq s \geq t_0$. The Markovian character is hence preserved by the time-dependent dynamical equation in (130).

A more general approach to the modeling of the dynamics of open quantum systems is based on the Nakajima-Zwanzig projection operator technique (Breuer and Petruccione, 2002; Nakajima, 1958; Zwanzig, 1960). According to that the dynamical equation is represented as follows:

$$\frac{d}{dt} \rho(t) = \int_{t_0}^t \mathcal{K}(t-u) \rho(u) du, \quad \rho(t_0) = \rho_0. \quad (133)$$

Here memory effects are modeled by the introduction of the *memory kernel* operator $\mathcal{K}(t)$. Hence, the rate of change of the state at time also depends on its history, and the Markovian setting (126) is reobtained when $\mathcal{K}(\tau) = 2\delta(\tau)\mathcal{L}$.

The dynamical map $\Lambda(t, t_0)$ associated to the non-Markovian evolution (133) is a solution of

$$\frac{d}{dt} \Lambda(t, t_0) = \int_{t_0}^t d\tau \mathcal{K}(t-\tau) \Lambda(\tau, t_0), \quad \Lambda(t_0, t_0) = \text{id}. \quad (134)$$

The dynamical map $\Lambda(t, t_0)$ appears to be a function of both t_0 and t_1 . However, one can notice that the dynamics of an open quantum system can be always understood as the reduced dynamics of an isolated one which includes the environment. Being the unitary dynamics of an isolated system homogeneous in time, it follows that, once the degree of freedom of the environment are taken into account, the dynamical map will be only a function of the difference ' $t - t_0$ ', that is, $\Lambda(t, t_0) \equiv \Lambda(t - t_0)$. Indeed, it has been proven in (Chruściński and Kossakowski, 2010) that any solution of (134) is also a solution of the time-dependent Markovian equation

$$\frac{d}{dt} \Lambda(t - t_0) = \mathcal{L}(t, t_0) \Lambda(t, t_0), \quad \Lambda(t_0, t_0) = \text{id}, \quad (135)$$

with a time-dependent Liouvillian defined by the logarithmic derivative of the dynamical map $\mathcal{L}(t - t_0) :=$

$\frac{d}{dt}\Lambda(t-t_0) \cdot \Lambda^{-1}(t-t_0)$. However, the explicit dependence of the generator on the initial time ‘ t_0 ’ implies that \mathcal{L} is effectively non-local in time. Finally, the formal solution of (135) reads

$$\Lambda(t, t_0) = \text{T exp} \left(\int_0^{t-t_0} \mathcal{L}(\tau) d\tau \right), \quad (136)$$

which shows that $\Lambda(t, t_0)$ is indeed homogeneous in time. However, contrary to (131), it does not satisfy the composition law, a fact which represents a signature for memory effects.

B. Legitimate memory kernels

One of the fundamental problems in the theory of non-Markovian master equations is to find those conditions on the memory kernel $\mathcal{K}(t)$ that ensure that the time evolution map $\Lambda(t, t_0)$ is completely positive and trace preserving, i.e., a quantum channel. Contrary to the Markovian case, a full characterization of legitimate memory kernels is still missing.

In (Chruściński and Kossakowski, 2012) a class of memory kernels giving rise to legitimate quantum dynamics (quantum channels) has been provided. The construction is based on a simple idea of normalization: starting from a family of completely positive maps satisfying a certain additional condition one is able to ‘normalize’ it in order to obtain legitimate, i.e., completely positive and trace preserving, quantum dynamics.

First of all, one notices that the non-Markovian master equation (134) is trace preserving iff $\text{Tr}[\mathcal{K}(t)\rho] = 0$ for any density operator ρ . Equivalently, this condition may be rewritten in the Heisenberg picture in terms of the dual of the memory kernel as $\mathcal{K}^*(t)I = 0$, where I denotes the identity operator.

Let us introduce an arbitrary family completely positive $N(t)$ which is differentiable and satisfies initial condition $N_0 = \text{id}$. Note that it may be represented as follows

$$N(t) = \text{id} - \int_0^t F(\tau) d\tau, \quad (137)$$

where $F(t) = -\dot{N}(t)$. It is possible to prove that if $F^*(t)I \geq 0$ there exists a family of completely positive maps $Q(t)$ satisfying

$$Q^*(t)I = F^*(t)I. \quad (138)$$

and such that the formula

$$\tilde{\mathcal{K}}(s) = [\tilde{Q}(s) - (\text{id} - s\tilde{N}(s))]\tilde{N}^{-1}(s), \quad (139)$$

where $\tilde{\cdot}$ denotes the Laplace transform, defines a legitimate memory kernel $\mathcal{K}(t)$.

As an example, if one starts from

$$N(t) = \left(1 - \int_0^t f(\tau) d\tau\right) \text{id},$$

where $f(t) \geq 0$, and $\int_0^\infty f(t) dt \leq 1$ to guarantee complete positivity, one obtains the memory kernel

$$\mathcal{K}(t) = \kappa(t)(\mathcal{A} - \text{id}),$$

with \mathcal{A} being an arbitrary quantum channel, and $\kappa(t)$ is defined through its Laplace transform,

$$\tilde{\kappa}(s) = \frac{s\tilde{f}(s)}{1 - \tilde{f}(s)}.$$

C. Markovian vs non-Markovian dynamics

The Gorini-Kossakowski-Sudarshan-Lindblad representation (126) characterizes the infinitesimal generator of Markovian dynamics. Nevertheless, a unique operative definition and quantifier of (non)Markovianity is still lacking, and several proposals have been put forward. Here we briefly review two of these proposals, respectively based on the notion of divisibility of a quantum channel, and distinguishability of quantum states.

A one-parameter family of CPTP maps $\Lambda(t, 0)$ is said to be infinitesimally divisible (Wolf and Cirac, 2008) if it can be written as a

$$\Lambda(t + \tau, 0) = \Lambda(t + \tau, t)\Lambda(t, 0), \quad (140)$$

with $\Lambda(t + \tau, t)$ a CPTP map for any $t, \tau > 0$. (Rivas, Huelga and Plenio, 2010) define a map to be (time-inhomogeneous) Markovian when it satisfies the composition law

$$\Lambda(s, t) = \Lambda(s, u)\Lambda(u, t), \quad (141)$$

for any $s \geq u \geq t$. This composition rule is indeed a characteristic trait of Markovian evolution: it is the quantum counterpart to the classic Chapman-Kolmogorov equation. Clearly, a family of CPTP maps satisfying (141) defines a one-parameter family of infinitesimally divisible quantum channels (140). A characterization of infinitesimally divisible quantum channels can be obtained in terms of the quantity

$$g(t) := \lim_{\epsilon \rightarrow 0+} \frac{\|\text{id}_d \otimes \Lambda(t + \epsilon, t)P_d^+\|_1 - 1}{\epsilon}, \quad (142)$$

where P_d^+ denotes the projector into a maximally entangled state dimension d and id_d is the identity map. It has been proven (Rivas, Huelga and Plenio, 2010) that $g(t) > 0$ if and only if the map $\Lambda(t, 0)$ is indivisible.

One fundamental property of CPTP maps is that they cannot increase the trace-distance, that is,

$$D(\Lambda(t, 0)(\rho_1), \Lambda(t, 0)(\rho_2)) \leq D(\rho_1, \rho_2), \quad (143)$$

for any pair of initial states ρ_1, ρ_2 . If a family of CPTP maps is infinitesimally divisible, the monotonicity of the trace distance holds true locally, that is,

$$\sigma(\rho_1, \rho_2; t) := \frac{d}{dt} D(\Lambda(t, 0)(\rho_1), \Lambda(t, 0)(\rho_2)) \leq 0. \quad (144)$$

This idea has been used to define a criterion of non-Markovianity (Breuer, Laine and Piilo, 2009). According to that the dynamical map $\Lambda(t, 0)$ is said to be non-Markovian if there exist a value of t such that $\sigma(\rho_1, \rho_2; t) > 0$, for some initial states ρ_1, ρ_2 . Physically, this implies a temporal increase in the distinguishability of the two quantum states, due to a backflow of information from the surrounding environment.

These two criteria allow one to define computable measure of non-Markovianity. A natural quantifier derived from the criterion of (Rivas, Huelga and Plenio, 2010) reads

$$\mathcal{N}_{\text{RHP}}(\Lambda) = \frac{\mathcal{I}}{\mathcal{I} + 1}, \quad (145)$$

where $\mathcal{I} = \int_0^\infty g(t)dt$. From the criterion of (Breuer, Laine and Piilo, 2009) one defines the non-Markovianity quantifier

$$\mathcal{N}_{\text{BLP}}(\Lambda) = \sup_{\rho_1, \rho_2} \int_{\sigma > 0} \sigma(\rho_1, \rho_2; t) dt. \quad (146)$$

(Chruściński, Kossakowski and Rivas, 2011) have pointed out that the relation these two criteria resembles that between separable and PPT states in entanglement theory. Indeed, any family of CPTP maps which is Markovian according to the first criterion is as well Markovian according to the second one, that is, $\mathcal{N}_{\text{RHP}}(\Lambda) = 0$ implies $\mathcal{N}_{\text{BLP}}(\Lambda) = 0$, while the converse is in general not true.

D. A solvable model

One of the few example of non Markovian dynamics that are exactly solvable for their communication capacities is a single qubit coupled to an environment of (non interacting) qubits giving rise to dephasing channel introduced by (Arshed, Toor and Lidar, 2010). Such a model is defined by a spin-star system of $N + 1$ spin-1/2, where the central system plays the role of the information carrier, while the other N spins model the surrounding environment. The Hamiltonian of the system is of the form $H = \sum_{n=0}^N \Omega_n \sigma_n^z$, where the label $n = 0$ refers to the central spin. Moreover, an interaction term is introduced between the central spin and the others: $H_I = \sigma_0^z \sum_{n=1}^N g_n \sigma_n^z$. The N surrounding spins are initially assumed to be in the thermal state associated to their own Hamiltonian. Since H commutes with H_I such an initial bath state is stationary.

Due to the form of the Hamiltonian the associated quantum channel, mapping the initial state of the central spin to its state at time t , is a dephasing channel for any value of $t \geq 0$. The quantum capacity of this quantum channel can be hence computed for all times, yielding

$$Q(t) = \frac{(1 + \pi_N) \log_2((1 + \pi_N) + (1 - \pi_N) \log_2(1 - \pi_N))}{2}, \quad (147)$$

where the function π_N , besides depending on time, strongly depends on the parameters entering the dynamical model. In turns, the behavior of the quantum capacity as function of time is as well strongly dependent on the couplings parameters and on the temperature of the bath. For generic values, recurrences in the capacity as function of time are of small amplitude and quickly vanish. On the other hand, for commensurable values of the parameters the quantum capacity is a periodic as function of time. This feature indicates the backflow of information from the environment to the central spin, a signature of non-Markovian dynamics. This is also related to the increased distinguishability of states pointed out by the BLP criterion.

IX. SUMMARY AND OUTLOOK

It has been seen that accounting for memory effects in quantum processes (channels) is a daunting task. However, the consideration of spatial and temporal memory effects is becoming increasingly pressing with the continuing miniaturization of devices and with increasing communication rates. For what concern the capacities evaluation, among the large variety of quantum memory channels the class that leaves hopes for an exact evaluation is that of forgetful channels. For this reason it is of utmost importance to derive general criteria to decide whether or not a given channel is forgetful. Beyond that it would be extremely interesting to establish when memory effects increase the capacity of a quantum channel. It is also worth noticing that the effects of correlations among errors are in close connection with the property of superadditivity of the minimum output entropy (Hastings, 2009). The possible memory induced enhancement of capacity of a quantum channel, looked through the dynamical memory model sketched in Sec. VII.A, can be seen as due to a sort of Zeno effect (Misra and Sudarshan, 1977). In fact, by frequently inserting information carriers through the channel one prevents the environment to come back to its stationary state after the passage of each of them, thus less affecting the carriers themselves. Whereas in the case of quantum channels arising from nonMarkovian dynamics like that of Sec. VIII.D, the increment of capacity can be explained by the backflow of information from environment to system.

Examples of quantum channels showing memory effects are abundant in quantum information processing.

An unmodulated spin chain has been proposed as a model for short distance quantum communication (Bose, 2003). In such a scheme, the state to be communicated over the channel is placed on one of the spins of the chain, propagates for a specific amount of time, and is then received at a distant spin of the chain. When viewed as a model for quantum communication, it is generally assumed that a reset of the spin chain occurs after each signal, for instance by applying an external magnetic field, resulting in a memoryless channel. However, a continu-

ous operation without resetting corresponds to a quantum channel with memory (Bayat *et al.*, 2008).

Another model of a quantum channel with memory is the so-called *one-atom maser* or *micromaser* (Benenti, D'arrigo and Falci, 2009). In such a device, excited atoms interact with the photon field inside a high-quality optical cavity. If the photons inside the cavity have sufficiently long lifetime, atoms entering the cavity will feel the effect of the preceding atoms, introducing correlations between consecutive signal states.

Furthermore, the propagation of the quantum electromagnetic field in linear dispersive media, including the free-space propagation and through linear optical systems, can be described by a quantum channel with memory (Giovannetti, *et al.*, 2004b; Lupo, *et al.*, 2011b; Shapiro, 2009), where wave diffraction introduces the memory effects.

Another source of correlated noise in the propagation of the electromagnetic field is due to atmospheric turbulence, whose effects on the signal propagation can be modeled as random changes of the channel's characteristics (Semenov and Vogel, 2009, 2010). Moreover, the decoherence induced by atmospheric turbulence introduces cross talks (Boyd *et al.*, 2011; Tyler and Boyd, 2009), i.e. intersymbol interference effects (see Sec. III.B.3), when information is encoded in the transverse degrees of freedom of the electromagnetic field, e.g. the orbital angular momentum.

Memory effects in quantum cryptography, that is to say private capacities, have been briefly addressed in (Ruggeri and Mancini, 2007b; Vasiel *et al.*, 2011) and needs further explorations. Then, an analysis of memory effects in other channel uses-configurations, like zero error channel capacity, channels with feedback, channels with unknown parameters and multiuser channels, should be pursued.

Non Markovian dynamics are relevant in several physical systems characterized by the interaction with a structured environment. Examples are in the framework of solid state physics, as quantum dots in photonic crystals (Madsen *et al.*, 2011; Vats, John and Busch, 2002), and in the soft matter framework as the exciton dynamics in the protein environment (Caruso *et al.*, 2009; Plenio and Huelga, 2008; Rebentrost, Chakraborty and Aspuru-Guzik, 2009; Thorwart *et al.*, 2009).

Finally, notice that correlated noise effects could be investigated not only in the framework of the quantum version of the second Shannon theorem, but also in the framework of the quantum version of the first one. In particular one could consider compression of correlated sources, or entanglement manipulation in the presence of correlated noise.

Acknowledgments

The authors would like to thank K. Banaszek, D. Chruscinski, N. Datta, M. B. Plenio, S. F. Huelga, O.

Pilyavets, S. Virmani, C. Macchiavello, N. J. Cerf, for useful discussions and for careful reading the manuscript.

S.M. acknowledges fruitful and enjoyable collaboration with Garry Andrew Bowen during the early stage development of the field of quantum memory channels.

The work of F.C. has been supported by a Marie Curie Intra-European Fellowship within the 7th European Community Framework Program and by MIUR-FIRB (Project No. RBFR10M3SB). The work of C.L. and S.M. has been supported by the European Commission, under the FET-Open grant agreement CORNER, number FP7-ICT-213681.

Appendix A: Tools for characterizing quantum channels

In order to quantify how much quantum information is preserved by some noisy channel Φ the notion of distance measures between quantum states needs to be introduced. Although there is a variety of such measures, here will focus on two widely used ones: *trace distance* and *fidelity*. The aim is to quantify how close two quantum states, represented by the density matrices ρ_1 and $\rho_2 \in \mathfrak{S}(\mathcal{H}_Q)$, are in the Hilbert space \mathcal{H}_Q . In this context, the trace distance of two quantum states ρ_1 and ρ_2 is defined as:

$$D(\rho_1, \rho_2) := \frac{1}{2} \|\rho_1 - \rho_2\|_1, \quad (\text{A1})$$

with $\|\Theta\|_1 := \text{Tr} \sqrt{\Theta^\dagger \Theta}$ being the *trace-norm*, or *norm-1* of the operator Θ . It is clearly a symmetric function of the two states, i.e., $D(\rho_1, \rho_2) = D(\rho_2, \rho_1)$, and vanishes if and only if $\rho_1 = \rho_2$. Furthermore, the trace distance is preserved under unitary transformation, i.e., $D(U\rho_1 U^\dagger, U\rho_2 U^\dagger) = D(\rho_1, \rho_2)$, but contractive under CPTP map Φ :

$$D(\Phi(\rho_1), \Phi(\rho_2)) \leq D(\rho_1, \rho_2). \quad (\text{A2})$$

Another important quantity, playing a key role in quantum communication theory, is the so-called fidelity F . Note that it does not define a distance in $\mathfrak{S}(\mathcal{H}_Q)$ since, for instance, it is equal to 1 when the two states are the same. Mathematically, the quantum fidelity between two states ρ_1 and ρ_2 , i.e., $F(\rho_1, \rho_2)$, is given by (Jozsa, 1994; Uhlmann, 1976):

$$F(\rho_1, \rho_2) := \left[\text{Tr} \left(\sqrt{\rho_1^{1/2} \rho_2 \rho_1^{1/2}} \right) \right]^2. \quad (\text{A3})$$

It is always in the range $[0, 1]$, equal to 1 if and only if $\rho_1 = \rho_2$, and symmetric, $F(\rho_1, \rho_2) = F(\rho_2, \rho_1)$. It vanishes for density matrices with orthogonal supports, e.g., it is vanishing for orthogonal pure states. Furthermore, it is invariant under the action of a unitary evolution, $F(U\rho_1 U^\dagger, U\rho_2 U^\dagger) = F(\rho_1, \rho_2)$, and monotonically increases under CPTP map,

$$F(\Phi(\rho_1), \Phi(\rho_2)) \geq F(\rho_1, \rho_2). \quad (\text{A4})$$

The *entanglement fidelity* (Schumacher, 1996) is another quantity which is often used to gauge how much quantum information is preserved under the action of a quantum channel Φ acting on a state $\rho \in \mathfrak{S}(\mathcal{H}_Q)$. To define the latter quantity, one introduces a reference system, described by the Hilbert space \mathcal{H}_R , isomorphic to the Hilbert space \mathcal{H}_Q . A purification of ρ is a pure state $|\psi_\rho\rangle \in \mathcal{H}_Q \otimes \mathcal{H}_R$ such that $\rho = \text{Tr}_R |\psi_\rho\rangle\langle\psi_\rho|$. The entanglement fidelity is then defined as

$$F(\rho, \Phi) := \langle\psi_\rho|\Phi(|\psi_\rho\rangle\langle\psi_\rho|)|\psi_\rho\rangle. \quad (\text{A5})$$

Given a set of Kraus operators $\{K_j\}_{j=1,\dots,d_E}$ for the quantum channel Φ , the entanglement fidelity can be written as follows:

$$F_e(\rho, \Phi) = \sum_{j=1}^{d_E} |\text{Tr}(\rho K_j)|^2. \quad (\text{A6})$$

In particular, if ρ is a maximally mixed state, that is, $\rho = I_Q / \dim \mathcal{H}_Q$, the entanglement fidelity reads

$$F_e(I_Q / \dim \mathcal{H}_Q, \Phi) = \frac{1}{(\dim \mathcal{H}_Q)^2} \sum_{j=1}^{d_E} |\text{Tr}(K_j)|^2. \quad (\text{A7})$$

Finally, the mathematical notion of norm of complete boundness, or cb-norm, is often used in the context of quantum channel theory. It is defined as:

$$\|\Phi\|_{cb} := \sup_n \|\Phi \otimes \text{id}_n\|_\infty, \quad (\text{A8})$$

where id_n denotes the identity map on $\mathfrak{S}(\mathbb{C}^n)$, and $\|\cdot\|_\infty$ is defined by

$$\|\Phi \otimes \text{id}_n\|_\infty := \sup_\rho \frac{\|\Phi \otimes \text{id}_n(\rho)\|_1}{\|\rho\|_1}. \quad (\text{A9})$$

This norm is also called diamond-norm, denoted $\|\cdot\|_\diamond$.

All these distance measures leads exactly to the same definitions of channel capacities (Kretschmann and Werner, 2004), as defined in the following in terms of fidelities and error correcting codes (see Secs. V, VI).

Another distance quantifier is the *quantum relative entropy*, which is introduced in the next subsection.

Appendix B: Entropic quantities

In the study of quantum communication, entropic quantities play a fundamental role in characterizing quantum channels in terms of their efficiency as communication lines (Barnum, Nielsen and Schumacher, 1998).

Let us denote by

$$S(\rho) := -\text{Tr}(\rho \log \rho), \quad (\text{B1})$$

the von Neumann entropy of a density operator ρ . Then, given a quantum channel Ψ and its input state ρ , there are three important entropic quantities related to the pair

(ρ, Φ) , namely, the entropy of the input state $S(\rho)$ (*input entropy*), the entropy of the output state $S(\Phi(\rho))$ (*output entropy*), and the *entropy of exchange* $S(\rho, \Phi)$. The entropy exchange is defined as

$$S(\rho, \Phi) := S[(\Phi \otimes \text{id})(|\psi_\rho\rangle\langle\psi_\rho|)], \quad (\text{B2})$$

that is, as the output entropy of the *dilated* channel $(\Phi \otimes \text{id})$ where the input state $|\psi_\rho\rangle$ is a purification of ρ . Such a quantity is independent of the choice of the purification. By construction the entropy of exchange can also be expressed as the output entropy of the complementary channel $\tilde{\Phi}$ of Φ , i.e., $S(\rho, \Phi) = S(\tilde{\Phi}(\rho))$. Moreover, these three quantities satisfy the triangle inequality

$$|S(\Phi(\rho)) - S(\rho, \Phi)| \leq S(\rho). \quad (\text{B3})$$

From these three entropies one can construct several information quantities. In analogy with classical information theory, one defines the *quantum mutual information* between the reference system R (which mirrors the input Q) and the output system Q' (Petz, 2008), as

$$\begin{aligned} \mathcal{I}(\rho, \Phi) &:= S(\rho'_R) + S(\rho'_Q) - S(\rho'_{RQ}) \\ &= S(\rho) + S(\Phi(\rho)) - S(\rho, \Phi). \end{aligned} \quad (\text{B4})$$

Notice that an important component of $\mathcal{I}(\rho, \Phi)$ is the *coherent information*

$$J(\rho, \Phi) := S(\Phi(\rho)) - S(\rho, \Phi). \quad (\text{B5})$$

The quantum mutual information \mathcal{I} satisfies the data-processing inequality:

$$\mathcal{I}(\rho, \Phi_2 \circ \Phi_1) \leq \min\{\mathcal{I}(\rho, \Phi_1), \mathcal{I}(\rho, \Phi_2)\}, \quad (\text{B6})$$

which in turn implies a similar inequality for the coherent information, i.e.,

$$J(\rho, \Phi_2 \circ \Phi_1) \leq J(\rho, \Phi_1), \quad (\text{B7})$$

where Φ_1 and Φ_2 are two generic quantum channels.

Finally, a distance-like entropic quantity is given by the *quantum relative entropy*. In particular, the quantum relative entropy of ρ_1 with respect to ρ_2 is defined as:

$$S(\rho_1 \| \rho_2) := \text{Tr}[\rho_1(\log \rho_1 - \log \rho_2)], \quad (\text{B8})$$

with ρ_1, ρ_2 being two generic density operators. This quantity is always nonnegative, equal to zero if and only if $\rho_1 = \rho_2$. A useful feature of this entropy measure is its monotonicity when a quantum channel Φ is applied to both states, i.e.,

$$S[\Phi(\rho_1) \| \Phi(\rho_2)] \leq S(\rho_1 \| \rho_2), \quad (\text{B9})$$

hence, two quantum states become less distinguishable after the action of a noisy quantum channel Φ (Lindblad, 1975).

Appendix C: Decomposition for non-anticipatory quantum channels

In this section, one provides an explicit derivation of the decomposition (17) based on a generalization of analysis presented in (Beckman *et al.*, 2001; Eggeling, Schlingemann and Werner, 2001; Kretschmann and Werner, 2005; Piani *et al.*, 2006).

Let then $\mathcal{F} = \{\Phi^{(n)}; n = 1, 2, \dots\}$ be a family of CPTP maps which describes a non-anticipatory quantum channel. Adopting the unitary representation (2), for each n integer one can define a unitary transformation $W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)}$ coupling the first n carriers to a common environment M which allows us to write

$$\Phi^{(n)}(\rho_Q^{(n)}) = \text{Tr}_M[W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)} (\rho_Q^{(n)} \otimes \Omega_M) [W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)}]^\dagger], \quad (\text{C1})$$

with $\Omega_M = |\Omega\rangle_M \langle \Omega|$ being a pure fixed state of M independent from the input state $\rho_Q^{(n)}$. In general the unitary couplings $W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)}$ can have a complicated dependence upon n : however since the channel is non-anticipatory they must obey the following rule

$$\begin{aligned} & \text{Tr}_{q_n, M}[W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)} (\rho_Q^{(n)} \otimes \Omega_M) [W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)}]^\dagger] \\ &= \text{Tr}_M[W_{q_{n-1}, \dots, q_1, M}^{(n-1)} (\rho_Q^{(n-1)} \otimes \Omega_M) [W_{q_{n-1}, \dots, q_1, M}^{(n-1)}]^\dagger], \end{aligned}$$

where $\rho_Q^{(n-1)} = \text{Tr}_{q_n}[\rho_Q^{(n)}]$ is the reduced density matrix of $\rho_Q^{(n)}$ associated with the first $n-1$ carriers. Applying this relation to a generic pure state $\rho_Q^{(n)}$ of the form $|\psi\rangle_{q_n} \otimes |\phi\rangle_{q_{n-1}, \dots, q_1}$ one notices that the vectors $W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)} |\psi\rangle_{q_n} \otimes |\phi\rangle_{q_{n-1}, \dots, q_1} \otimes |\Omega\rangle_M$ and $W_{q_{n-1}, \dots, q_1, M}^{(n-1)} |\psi\rangle_{q_n} \otimes |\phi\rangle_{q_{n-1}, \dots, q_1} \otimes |\Omega\rangle_M$ are both purification of the state $\Phi^{(n-1)}(\rho_Q^{(n-1)})$ constructed on the ancillary system formed by q_n and M . Therefore there must exist a unitary transformation $U_{q_n M}$ acting on the latter which satisfies the identity

$$W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)} |\Omega\rangle_M = U_{q_n M} W_{q_{n-1}, \dots, q_1, M}^{(n-1)} |\Omega\rangle_M. \quad (\text{C2})$$

Iterating this n times yields

$$W_{q_n, q_{n-1}, \dots, q_1, M}^{(n)} |\Omega\rangle_M = U_{q_n M} U_{q_{n-1} M} \cdots U_{q_1 M} |\Omega\rangle_M \quad (\text{C3})$$

which replaced into Eq. (C1) implies Eq. (17).

As a final remark, one notices that the above construction is always possible if one assumes M to have dimension $(\dim \mathcal{H}_q)^n$: in particular this implies that one can take M to be a composite system formed by n subsystem m_1, m_2, \dots, m_n initialized into the pure factorized state $|\Omega\rangle_M = |0\rangle_{m_1} \otimes \cdots \otimes |0\rangle_{m_n}$. Under this assumption the operator $W_{q_{n-1}, \dots, q_1, M}^{(n-1)}$ can be chosen to act non trivially only on the first $n-1$ environmental components: consequently the $U_{q_j M}$ can be chosen to couple q_j only with the corresponding j -th components of M . In this case the associated unitary representation (2) corresponds to the one of Fig. 1 b).

Appendix D: C^* - and Quasi-Local Algebras

C^* -algebras are axiomatic generalizations of notions like continuous functions over compact spaces or bounded operators over Hilbert spaces. They are realized by supplying the algebraic structure of these elements (resulting from the possibility to add and multiply them) by an adjoint operation and a norm. An easy introduction to C^* -algebras is given in (Kadison and Ringrose, 1997).

Let \mathcal{A} be an algebra over the field \mathbb{C} , i.e. a vector space over \mathbb{C} equipped with a distributive and associative product. Introduce an adjoint operation $\dagger : \mathcal{A} \rightarrow \mathcal{A}$ which is an anti-linear map $(\lambda a + \mu a')^\dagger = \bar{\lambda} a^\dagger + \bar{\mu} a'^\dagger$, with $a^{\dagger\dagger} = a$ and $(aa')^\dagger = a'^\dagger a^\dagger$ for $a, a' \in \mathcal{A}, \lambda, \mu \in \mathbb{C}$. Then, the algebra \mathcal{A} over \mathbb{C} equipped with \dagger is called $*\text{-algebra}$. In turn, such an algebra will be termed $C^*\text{-algebra}$ if there is a norm $\|\cdot\| : \mathcal{A} \rightarrow [0, \infty)$ such that \mathcal{A} is complete with respect to $\|\cdot\|$, $\|aa'\| \leq \|a\| \cdot \|a'\|$ for all $a, a' \in \mathcal{A}$ and $\|a^\dagger a\| = \|a\|^2$ for all $a \in \mathcal{A}$. Furthermore, if a C^* -algebra \mathcal{A} possesses an (identity) element $\mathbf{1} \in \mathcal{A}$, such that $\mathbf{1}a = a$ for all $a \in \mathcal{A}$, is called *unital*. Then on a unital C^* -algebra \mathcal{A} is possible to define a state as a linear functional $\psi : \mathcal{A} \rightarrow \mathbb{C}$ which is positive, i.e. $\psi(a^*a) \geq 0$, and normalized, i.e. $\psi(\mathbf{1}) = 1$.

By considering the algebra $\mathcal{A} = \mathcal{B}(\mathcal{H})$ on a (finite-dimensional) Hilbert space \mathcal{H} together with the Hilbert-Schmidt product $\langle a, b \rangle = \text{tr}(a^\dagger b)$, $a, b \in \mathcal{A}$ and the Riesz theorem, it can be easily seen that each state ψ on \mathcal{A} is represented by a unique density operator $\rho \in \mathcal{A}$, that is $\psi(a) = \text{tr}(\rho a)$ for all $a \in \mathcal{A}$. Notice that for infinite dimensional systems there may be states not representable by density operators in this sense.

A $*\text{-automorphism}$ of a C^* -algebra \mathcal{A} is a one-to-one linear map $T : \mathcal{A} \rightarrow \mathcal{A}$ with $T(a^\dagger) = (T(a))^\dagger$ and $T(aa') = T(a)T(a')$ for all $a, a' \in \mathcal{A}$. Any $*\text{-automorphism}$ induces a family $\{T_z\}_{z \in \mathbb{Z}}$ of $*\text{-automorphisms}$, with $T_{z_1+z_2} = T_{z_1} \circ T_{z_2}$ and $T_0 = \text{id}_{\mathcal{A}}$, given by $T_z := T^z$, $z \in \mathbb{Z}$.

A state ψ on \mathcal{A} is said to be \mathbb{Z} -invariant, or T -invariant, if $\psi \circ T = \psi$ holds true. Obviously $\psi \circ T_z = \psi$ for all $z \in \mathbb{Z}$. Moreover, the set of \mathbb{Z} -invariant states is convex. Now comes the definition of *ergodic* state ψ as an extreme point in the convex set of \mathbb{Z} -invariant states.

Quasi-local algebras are the proper mathematical tools to describe infinitely extended quantum lattice systems (see (Simon, 1993) for a detailed review on the subject). Let us consider a finite-dimensional C^* -algebra \mathcal{A} and suppose to attach a copy \mathcal{A}_n of \mathcal{A} to each $n \in \mathbb{Z}$. Furthermore, let $\Lambda \subset \mathbb{Z}$ be a finite set and define $\mathcal{A}^\Lambda := \bigotimes_{n \in \Lambda} \mathcal{A}_n$. Then, \mathcal{A}^Λ is called the algebra of observables belonging to all sites in Λ . Whenever $\Lambda \subset \Lambda' \subset \mathbb{Z}$ tensoring with identity $\mathbf{1}_{\Lambda' \setminus \Lambda}$ in $\mathcal{A}^{\Lambda \setminus \Lambda}$ makes \mathcal{A}^Λ sub-algebra of $\mathcal{A}^{\Lambda'}$ with an embedding $\mathcal{A}^\Lambda \ni a \mapsto a \otimes \mathbf{1}_{\Lambda' \setminus \Lambda} \in \mathcal{A}^{\Lambda'}$. In this way the product of operators, as well as other algebraic operations, become well defined in the larger algebra $\mathcal{A}^{\Lambda \cup \Lambda'}$. Thus, since the identity $\mathbf{1}_{\Lambda \setminus \Lambda}$ does not change the norm, one can define the normed $*\text{-algebra}$ of *local observables* $\mathcal{A}^{\text{loc}} := \bigcup_{\Lambda \subset \mathbb{Z}; |\Lambda| < \infty} \mathcal{A}^\Lambda$. Its norm-

completion

$$\mathcal{A}^{\mathbb{Z}} := \overline{\mathcal{A}^{loc}} \quad (D1)$$

is then called the *quasi-local C^* -algebra*.

The algebra \mathcal{A}^{Λ} can be interpreted as the algebra of physical observables for a subsystem localized in the region $\Lambda \subset \mathbb{Z}$. The quasi-local algebra $\mathcal{A}^{\mathbb{Z}}$ then corresponds to the extended algebra of observables on the infinite one-dimensional lattice \mathbb{Z} . In the quasi-local algebra each observable can be approximated uniformly by local ones,

Any state ψ on $\mathcal{A}^{\mathbb{Z}}$ induces a family of states $\{\psi^{\Lambda}\}_{\Lambda \subset \mathbb{Z}: |\Lambda| < \infty}$ on \mathcal{A}^{Λ} and $\Lambda \subset \Lambda'$ implies that the restriction of $\psi^{\Lambda'}$ to \mathcal{A}^{Λ} is equal to ψ^{Λ} .

A shift $T: \mathcal{A}^{\mathbb{Z}} \rightarrow \mathcal{A}^{\mathbb{Z}}$ on $\mathcal{A}^{\mathbb{Z}}$ is induced by

$$\mathcal{A}^{\Lambda} \ni a \simeq a \otimes \mathbf{1}_{\Lambda} \mapsto T(a) := \mathbf{1}_{\Lambda} \otimes a \simeq a \in \mathcal{A}^{\Lambda+1}.$$

Notice that the extension of T to the quasi-local algebra $\mathcal{A}^{\mathbb{Z}}$ is a $*$ -automorphism. Moving from the action of the shift T , it is possible to introduce the notion of *stationary* state ψ on $\mathcal{A}^{\mathbb{Z}}$ when $\psi \circ T = \psi$ holds true. The set of stationary states on $\mathcal{A}^{\mathbb{Z}}$ turns out to be convex. Then a state ψ on $\mathcal{A}^{\mathbb{Z}}$ is called *ergodic* if it is extremal on this set.

References

- Adami, C. and Cerf, N. J., 1997, Phys. Rev. A **56**, 3470.
- Adesso, G. and Ericsson, M., 2006, Phys. Rev. A **74**, 030305(R).
- Ahlsvede, R., 1968, Z. Wahrscheinlichkeitstheor. Verwandte Geb. **11**, 61.
- Alicki, R. and Lendi, K., 1987, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics (Springer-Verlag, Berlin).
- Amosov, G. G. and Mancini, S., 2009, Quant. Inf. & Comp. **7**, 594.
- Arshed, N., Toor, A. H. and Lidar, D. A., 2010, Phys. Rev. A **81**, 062353.
- Ball, J., Dragan, A. and Banaszek, K., 2004, Phys. Rev. A **69**, 042324.
- Ball, J. and Banaszek, K., 2005, Open Sys. & Inf. Dynamics **12**, 121.
- Ban, M., Sasaki, M. and Takeoka, M., 2002, J. Phys. A **35**, L401.
- Banaszek, K., Dragan, A., Wasilewski, W. and Radzewicz, C., 2004, Phys. Rev. Lett. **92**, 257901.
- Barreto Lemos, G. and Benenti, G., 2010, Phys. Rev. A **81**, 062331.
- Barnum, H., Nielsen, M. A. and Schumacher, B., 1998, Phys. Rev. A **57**, 4153.
- Bartlett, S. D., Rudolph, T. and Spekkens, R. W., 2003, Phys. Rev. Lett. **91**, 027901.
- Bayat, A., Burgarth, D., Mancini, S. and Bose, S., 2008, Phys. Rev. A **77**, 050306(R).
- Beckman, D., Gottesman, D., Nielsen, M. A. and Preskill, J., 2001, Phys. Rev. A **64**, 052309.
- Ben-Aroya, A. and Ta-Shma, A., "Approximate quantum error correction for correlated noise", arXiv:0909.1466.
- Benenti, G., D'Arrigo, A. and Falci, G., 2009, Phys. Rev. Lett. **103**, 020502.
- Bengtsson, I., Zyczkowski, K., 2006, *Geometry of Quantum States*, (Cambridge, Cambridge University Press).
- Bennett, C. H., Brassard, G., Jozsa, R., Peres, A. and Wootters, W., 1993, Phys. Rev. Lett. **70**, 1895.
- Bennett, C. H., DiVincenzo, D. P. and Smolin, J. A., 1997, Phys. Rev. Lett. **78**, 3217.
- Bennett, C. H. and Shor, P. W., 1998, IEEE Trans. Inf. Theory **44**, 2724.
- Bennett, C. H., Shor, P. W., Smolin, J. A. and Thapliyal, A. V., 1999, Phys. Rev. Lett. **83**, 3081.
- Bennett, C. H., Shor, P. W., Smolin, J. A. and Thapliyal, A. V., 2002, IEEE Trans. Inform. Theory **48**, 2637.
- Bjelaković, I. and Boche, H., 2008, IEEE Trans. Inf. Th. **54**, 723.
- Bjelaković, I. and Boche, H., 2009, IEEE Trans. Inf. Th. **57**, 3360.
- Bjelaković, I., Boche, H. and Nötzel, J., 2009, Comm. Math. Phys. **292**, 55.
- Blackwell, D., Breiman, L. and Thomasian, A. J., 1958, Ann. Math. Stat. **29**, 1209.
- Bose, S., 2003, Phys. Rev. Lett. **91**, 207901.
- Bowen, G. and Bose, S., 2004, Phys. Rev. Lett. **87**, 267901.
- Bowen, G. and Mancini, S., 2004, Phys. Rev. A **69**, 012306.
- Bowen, G., Devetak, I. and Mancini, S., 2005, Phys. Rev. A **71**, 034310.
- Boyd, R. W., Rodenburg, B., Mirhosseini, M. and Barnett, S. M., 2011, Opt. Express **19**, 18310.
- Brandão, F. G. S. L., Oppenheim, J. and Strelchuk, S., 2012, Phys. Rev. Lett. **108**, 040501.
- Braunstein, S. L. and Kimble, H. J., 1998, Phys. Rev. Lett. **80**, 469.
- Braunstein, S. L. and van Loock, P., 2005, Rev. Mod. Phys. **77**, 513.
- Breuer, H. P. and Petruccione, F., 2002, *The Theory of Open Quantum Systems*, (Oxford, Oxford University Press).
- Breuer, H. P., Laine, E. M. and Piilo, J., 2009, Phys. Rev. Lett. **103**, 210401.
- Cafaro, C. and Mancini, S., 2010, Phys. Lett. A **374**, 2688.
- Cafaro, C. and Mancini, S., 2010, Phys. Rev. A **82**, 012306.
- Cafaro, C., L'Innocente, S., Lupo, C. and Mancini, S., 2011, Open Sys. & Inf. Dynamics **18**, 1.
- Cafaro, C. and Mancini, S., 2011, Int. J. Quant. Inf. **9**, 309.
- Calderbank, A.R. and Shor, P., 1996, Phys. Rev. A **54**, 1096.
- Calderbank, A. R., Rains, E. M., Shor, P. W. and Sloane, N. J. A., 1997, Phys. Rev. Lett. **78**, 405.
- Caruso, F. and Giovannetti, V., 2006, Phys. Rev. A **74**, 062307.
- Caruso, F., Giovannetti, V. and Holevo, A. S., 2006, New J. Phys. **8**, 310.
- Caruso, F. and Giovannetti, V., 2007, Phys. Rev. A **76**, 042331.
- Caruso, F., Eisert, J., Giovannetti, V. and Holevo, A. S., 2008, New J. Phys. **10**, 083030.
- Caruso, F., Giovannetti, V., Macchiavello, C. and Ruskai, M. B., 2008, Phys. Rev. A **77**, 052323.
- Caruso, F., Chin, A. W., Datta, A., Huelga, S. F. and Plenio, M. B., 2009, J. Chem. Phys. **131**, 105106.
- Caruso, F., Giovannetti, V. and Palma, M., 2010, Phys. Rev. Lett. **104**, 020503.
- Caves, C. M. and Drummond, P. D., 1994, Rev. Mod. Phys. **66**, 481.
- Cerf, N., Clavareau, J., Macchiavello, C. and Roland, J., 2005,

- Phys. Rev. A **72**, 042330.
- Cerf, N., Clavareau, J., Macchiavello, C. and Roland, J., 2005, *Int. J. Quantum Inf.* **4**, 439.
- Charmichael, H. J., 1993, *Phys. Rev. Lett.* **70**, 2273.
- Chau, H. F., 1998, *Phys. Rev. A* **58**, 905.
- Chau, H. F., 1999, *Phys. Rev. A* **60**, 1966.
- Chiribella, G., Dall'Arno, M., D'Ariano, G. M., Macchiavello, C. and Perinotti, P., 2011, *Phys. Rev. A* **83**, 052305.
- Choi, M. D., 1975, *Linear Algebra Appl.* **10**, 285.
- Chruściński, D. and Kossakowski, A., 2010, *Phys. Rev. Lett.* **104**, 070406.
- Chruściński, D., Kossakowski, A., and Rivas, A., 2011, *Phys. Rev. A* **83**, 052128.
- Chruściński, D. and Kossakowski, A., 2012, *Europhys. Lett.* **97**, 20005.
- Clemens, J. P., Siddiqui, S., and Gea-Banacloche, J., 2004, *Phys. Rev. A* **69**, 062313.
- Cover, T. M., and Thomas, J. A., 1968, *Elements of Information Theory*, (New York, Wiley).
- D'Arrigo, A., Benenti, G., and Falci, G., 2007, *New J. Phys.* **9**, 310.
- D'Arrigo, A., De Leo, E., Benenti, G., and Falci, G., 2008, *Int. J. Quant. Inf.* **6**, 651.
- Daffer, S., Wódkiewicz, K., and McIver, J. K., 2003, *Phys. Rev. A* **67**, 062312.
- Daems, D., 2007, *Phys. Rev. A* **76**, 012310.
- Daffer, S., Wódkiewicz, K., Cresser, J. D., and McIver, J. K., 2004, *Phys. Rev. A* **70**, 010304(R).
- Datta, N., and Dorlas, T. C., 2007, *J. Phys. A* **40**, 8147.
- Datta, N., Suhov, Y., and Dorlas, T. C., 2008, *Quant. Inf. Process.* **7**, 251.
- Datta, N. and Dorlas, T. C., 2009, *J. Stat. Phys.* **134**, 1173.
- Demkowicz-Dobrzanski, R., Kolenderski, P. and Banaszek, K., 2007, *Phys. Rev. A* **76**, 022302.
- Devetak, I., 2005, *IEEE Trans. Inf. Th.* **51**, 44.
- Devetak, I. and Shor, P.W., 2005, *Commun. Math. Phys.* **256**, 287.
- Dobrushin, R. L., 1963, *Am. Math. Soc. Transl.* **33**, 323.
- Eggeling, T., Schlingemann, D. M., and Werner, R. F., 2001, *Europhys. Lett.* **57**, 782.
- Eisert, J., and Plenio, M. B., 2003, *Int. J. Quant. Inf.* **1**, 479.
- Ekert, E., and Macchiavello, C., 1996, *Phys. Rev. Lett.* **77**, 2585.
- Feinstein, A., 1954, *IEEE Trans. Inf. Th.* **4**, 2.
- Feinstein, A., 1959, *Inform. Contr.* **2**, 25.
- Ferraro, A., Olivares, S., and Paris, M. G. A., 2005, *Gaussian states in continuous variable quantum information*, (Napoli, Bibliopolis).
- Gaitan, F., 2008, *Quantum Error Correction and Fault Tolerant Quantum Computing*, (Boca Raton, CRC Press).
- Gallager, R. G., 1968, *Information Theory and Reliable Communication*, (New York, Wiley).
- Gardiner, C. W., 1993, *Phys. Rev. Lett.* **70**, 2269.
- Gardiner, C. W. and Parkins, A. S. 1994, *Phys. Rev. A* **50**, 1792.
- Giovannetti, V., Lloyd, S., Maccone, L., and Shor, P. W., 2003, *Phys. Rev. A* **68**, 062323.
- Giovannetti, V., Guha, S., Lloyd, S., Maccone, L., Shapiro, J. H., and Yuen, H. P., 2004, *Phys. Rev. Lett.* **92**, 027902.
- Giovannetti, V., Guha, S., Lloyd, S., Maccone, L., Shapiro, J. H., Yen, B. J., and Yuen, H. P., 2004, *Quantum Inf. Comput.* **4**, 489.
- Giovannetti, V., and Fazio, R., 2005, *Phys. Rev. A* **71**, 032314.
- Giovannetti, V., 2005, *J. Phys. A* **38**, 10989.
- Giovannetti, V., and Mancini, S., 2005, *Phys. Rev. A* **71**, 062304.
- Giovannetti, V., Burgarth, D., and Mancini, S., 2009, *Phys. Rev. A* **79**, 012311.
- Giovannetti, V., Holevo, A. S., Lloyd, S., and Maccone, L., 2010, *J. Phys. A* **43**, 415305.
- Giovannetti, V. and Palma, M., 2012, *Phys. Rev. Lett.* **108**, 040401.
- Gorini, V., Kossakowski, A. and Sudarshan, E. C. G., 1976, *J. Math. Phys.* **17**, 821.
- Gottesman, D., 1997, *Stabilizer Codes and Quantum Error Correction*, PhD Thesis California, Institute of Technology, Pasadena Ca.
- Gottesman, D. "An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation", arXiv:quant-ph/0904.2557 (2009).
- Grassl, M. and Roetteler, M., 2006, *IEEE International Symposium on Information Theory*, pp.pp. 1109-1113.
- Gray, R. M., 1972, *IEEE Trans. Inf. Th.* **18**, 725.
- Gray, R. M., 2006, *Toeplitz and Circulant Matrices: A Review*, (Norwell, Now Publisher).
- Gütschow, J. "Representation of Convolutional Stabilizer Codes as Clifford Memory Channels", CORNER, EU project number FP7-ICT-213681 (internal report).
- Gütschow, J., Uphoff, S., Werner, R. F. and Zimborás, Z., 2010, *J. Math. Phys.* **51**, 015203.
- Hamada, M., 2002, *J. Math. Phys.* **43**, 4382.
- Hastings, M. B., 2009, *Nat. Phys.* **5**, 255.
- Hayashi, M., and Nagaoka, H., 2003, *IEEE Trans. Inf. Th.* **49**, 1753.
- Helstrom, C. W., 1976, *Quantum Detection and Estimation Theory*, (London, Academic Press).
- Hiroshima, T., 2006, *Phys. Rev. A* **73**, 012330.
- Holevo, A. S., 1998, *IEEE Trans. Inf. Theory* **44**, 269.
- Holevo, A. S. 2011, *Probabilistic and Statistical Aspects of Quantum Theory*, (Pisa: Edizioni della Normale).
- Holevo, A. S., Shohma, M., and Hirota, O., 1999, *Phys. Rev. A* **59**, 1820.
- Holevo, A. S., and Werner, R. F., 2001, *Phys. Rev. A* **63**, 032312.
- Holevo, A. S., 2007, *Probl. Inf. Transm.* **43**, 1.
- Holevo, A. S., Giovannetti, V., 2012, *Rep. Prog. Phys.*
- Piani, M., Horodecki, M., Horodecki, P., and Horodecki, R., 1996, *Phys. Lett. A* **223**, 1.
- Horodecki, M., Shor, P. and Ruskai, M.B., 2003, *Reviews in Mathematical Physics* **15**, 629-641.
- Houshmand, M., Hosseini-Khayat, S. and Wilde, M. M., "Minimal-memory, non-catastrophic, polynomial-depth quantum convolutional encoders, arXiv:1105.0649 (2011).
- Hu, Y., 2007, *Phys. Rev. A* **75**, 012314.
- Hwang, W. Y., Ahn, D., and Hwang, S. W., 2001, *Phys. Rev. A* **63**, 022303.
- Jacobs, K., 1962, *Colloquium on Combinatorial Methods in Probabilistic Theory*, (Aarhus).
- Jamiołkowski, A., 1972, *Rep. Math. Phys.* **3**, 275.
- Jozsa, R., 1994, *J. Mod. Opt.* **41**, 2315.
- Kawataba, S., 2000, *J. Phys. Soc. Jap.* **69**, 3540.
- Khinchin, A. I., 1957, *Mathematical Foundations of Information Theory*, (Dover Publications).
- Kadison, R. V. and Ringrose, J. R., 1997, *Fundamentals of the Theory of Operator Algebras Vol. II*, (American Mathematical Society).
- Karimipour, V. and Memarzadeh, L., 2006, *Phys. Rev. A* **74**,

- 032332.
- Karimipour, V. and Memarzadeh, L., 2006, Phys. Rev. A **74**, 062311.
- Karpov, E., Daems, D. and Cerf, N. J., 2006, Phys. Rev. A **74**, 032320.
- Karpov, E., Daems, D. and Cerf, N. J., 2006, Open Sys. & Inf. Dynamics **13**, 363.
- Keyl, M., 2002, Phys. Rep. **369**, 431-548.
- Khinchin, A. I., 1956, Usp. Math. Nauk. **11**, 17.
- King, C., 2003, IEEE Trans. Inf. Th. **49**, 221-229.
- Klesse, R. and Frank, S., 2005, Phys. Rev. Lett. **95**, 230503.
- Knill, E., et al., "Introduction to Quantum Error Correction", arXiv:quant-ph/020717 (2002).
- Knill, E., and Laflamme, R., 1997, Phys. Rev. A **55**, 900.
- Kraus, K., 1971, Ann. Phys. **64**, 311.
- Kretschmann, D., and Werner, R. F., 2004, New J. Phys. **6**, 26.
- Kretschmann, D., and Werner, R. F., 2005, Phys. Rev. A **72**, 062323.
- Kribs, D. W., Pasieka, A. and Zyczkowski, K., 2008, Open Sys. & Inf. Dynamics **15**, 329.
- Laflamme, R., Miquel, C., Paz, J. P., and Zurek, W. H., 1996, Phys. Rev. Lett. **77**, 198.
- Lidar, D. A., and Whaley, K. B., 2003 "Decoherence-Free Subspaces and Subsystems", in "Irreversible Quantum Dynamics", F. Benatti and R. Floreanini (Eds.), pp. 83-120 (Springer Lecture Notes in Physics vol. 622, Berlin, 2003), arXiv:quant-ph/0301032 (2003).
- Lidar, D. A., Bacon, D. and Whaley, K. B., 1999, Phys. Rev. Lett. **82**, 4556.
- Lindblad, G., 1975, Comm. Math. Phys. **40**, 147-151.
- Lindblad, G., 1976, Comm. Math. Phys. **48**, 119.
- Lloyd, S., 1997, Phys. Rev. A **55**, 1613.
- Lupo, C., Memarzadeh, L., and Mancini, S., 2009, Phys. Rev. A **80**, 042328.
- Lupo, C., Pilyavets, O. and Mancini, S., 2009, New J. of Phys. **11**, 063023.
- Lupo, C., Giovannetti, V., and Mancini, S., 2010, Phys. Rev. Lett. **104**, 030501.
- Lupo, C., Giovannetti, V., and Mancini, S., 2010, Phys. Rev. A **82**, 035007.
- Lupo, C., and Mancini, S., 2010, Phys. Rev. A **81**, 052314.
- Lupo, C., Pirandola, S., Aniello, P., and Mancini, S., 2011, Phys. Scr. T **143**, 014016.
- Lupo, C., Giovannetti, V., Pirandola, S., Mancini, S., and Lloyd, S., 2011, Phys. Rev. A **84**, 010303(R).
- Lupo, C., Giovannetti, V., Pirandola, S., Mancini, S., and Lloyd, S., 2012, Phys. Rev. A **85**, 062314.
- Macchiavello, C., and Palma, G. M., 2002, Phys. Rev. A **65**, 050301(R).
- Macchiavello, C., Palma, G. M., and Virmani, S., 2004, Phys. Rev. A **69**, 010303(R).
- Madsen, K. H., Ates, S., Lund-Hansen, T., Löffler, A., Reitzenstein, S., Forchel, A. and Lodahl, P., 2011, Phys. Rev. Lett. **106**, 233601.
- McMillan, B., 1953, Ann. Math. Stat.. **24**, 196.
- Memarzadeh, L., and Mancini, S., 2010, Eur. J. Phys. D **56**, 291.
- Misra, B. and Sudarshan, E. C. G., 1977, J. Math. Phys. **18**, 756.
- Nakajima, S., 1958, Prog. Theor. Phys. **20**, 948.
- Nielsen, M. A., "The entanglement fidelity and quantum error correction", arXiv: quant-ph/9606012 (1996).
- Nielsen, M. A. and Chuang, I. L., 2000, *Quantum Computation and Quantum Information*, (Cambridge, Cambridge University Press).
- Norris, J. R., 1997, *Markov Chains*, (Cambridge, Cambridge University Press).
- Novais, E., and Baranger, H. U., 2006, Phys. Rev. Lett. **97**, 040501.
- Ollivier, H., and Tillich, J.-P., 2003, Phys. Rev. Lett. **91**, 177902.
- Ollivier, H., and Tillich, J.-P., "Quantum convolutional codes: fundamentals", arXiv: quant-ph/0401134 (2004).
- Peres, A., 1996, Phys. Rev. Lett. **77**, 1413.
- Peterson, W. W., and Weldon, E. J., 1972, *Error Correcting Codes*, (Cambridge, MIT Press).
- Petz, D., 2008, *Quantum Information Theory and Quantum Statistics*, (Berlin, Springer).
- Piani, M., Horodecki, M., Horodecki, P., and Horodecki, R., 2006, Phys. Rev. A **74**, 012305.
- Pilyavets, O., Zborovskii, V., and Mancini, S., 2008, Phys. Rev. A **77**, 052324.
- Pilyavets, O., Lupo, C., and Mancini, S., 2009, arxiv 0907.1532.
- Plenio, M. B., and Huelga, S.F., 2008, New J. Phys. **99**, 113019.
- Plenio, M. B., and Virmani, S., 2007, Phys. Rev. Lett. **99**, 120504.
- Plenio, M. B., and Virmani, S., 2008, New J. Phys. **10**, 043032.
- Popescu, S., and Rohrlich, D., 2001, Found. Phys. **24**, 379.
- Rains, E.M., 2001, IEEE Trans. Inf. Theory **47**, 2921-2933.
- Rebentrost, P., Chakraborty, R. and Aspuru-Guzik, A., 2009, J. Chem. Phys. **131**, 184102.
- Rivas, A., Huelga, S. F., and Plenio, M. B., 2010, Phys. Rev. Lett. **105**, 050403.
- Ruggeri, G., Soliani, G., Giovannetti, V., and Mancini, S., 2005, Europhys. Lett. **70**, 719.
- Ruggeri, G., and Mancini, S., 2007a, Quant. Inf. & Comp. **7**, 265.
- Ruggeri, G., and Mancini, S., 2007b, Phys. Lett. A **362**, 340.
- Ruskai, M. B., 2003, Rev. Math. Phys. **15**, 643.
- Scarani, V., Ziman, M., Stelmachovic, P., Gisin, N. and Buzek, V., 2002, Phys. Rev. Lett. **88**, 097905.
- Schäfer, J., Karpov, E., and Cerf, N. J., 2009, Phys. Rev. A **80**, 062313.
- Schäfer, J., Karpov, E., and Cerf, N. J., 2010, Proc. of SPIE **7727**, 77270J.
- Schäfer, J., Karpov, E., and Cerf, N. J., 2011, Phys. Rev. A **84**, 032318.
- Schäfer, J., Karpov, E., and Cerf, N. J., 2012, Phys. Rev. A **85**, 012322.
- Schuch, N., Cirac, J. I. and Wolf, M. M., 2008 "Quantum Information and Many Body Quantum Systems", M. Ericsson and S. Montangero (Eds.), pp. 129-142 (Edizioni della Normale, Pisa, 2008), arXiv:quant-ph/0509166 (2005).
- Schumacher, B., 1996, Phys. Rev. A **54**, 2614.
- Schumacher, B. and Westmoreland, W. D., 1997, Phys. Rev. A **56**, 131.
- Semenov, A. A. and Vogel, W., 2009, Phys. Rev. A **80**, 021802.
- Semenov, A. A. and Vogel, W., 2010, Phys. Rev. A **81**, 023835.
- Shabani, A., 2008, Phys. Rev. A **77**, 022323.
- Shadman, Z., Kampermann, K., Bruss, D. and Macchiavello, C., 2011, Phys. Rev. A **84**, 042309.
- Shannon, C. E., 1948, Bell Syst. Tech. J. **27**, 379,623.

- Shannon, C. E., 1949, Proc. IRE **37**, 10.
- Shannon, C. E., 1957, Inform. Contr. **1**, 6.
- Shapiro, J. H., 2009, IEEE J. Sel. Top. Quantum Electron. **15**, 1547.
- Shor, P. W., 2002, *The quantum channel capacity and coherent information*, Lecture notes, MSRI Workshop on Quantum Computation.
- Simon, B., 1993, *The Statistical Mechanics of Lattice Gases*, (Princeton University Press).
- Smith, G., and Yard, J., 2010, Science **321**, 1812.
- Steane, A. M., 1996, Proc. R. Soc. London A **452**, 2551.
- Terhal, B. M. and DiVincenzo, D. P., 2000, Phys. Rev. A **61**, 022301.
- Thorwart, M., Eckel, J., Reina, J. H., Nalbach, P. and Weiss, S., 2009, Chem. Phys. Lett. **478**, 234.
- Tsaregradsky, I. P., 1958, Theory Probab. Applications **3**, 84.
- Tyler, G. A. and Boyd, R. W., 2009, Opt. Lett. **34**, 142.
- Uhlmann, A., 1976, Rep. Math. Phys. **9**, 273.
- Vaidman, L., 1994, Phys. Rev. A **49**, 1473.
- Vasile, R., Olivares, S., Paris, M. G. A. and Maniscalco, S., 2011, Phys. Rev. A **83**, 042321.
- Vatan, F., Roychowdhury, V. P., and Anantram, M. P., "Spatially Correlated Qubit Errors and Burst-Correcting Quantum Codes", arXiv:quant-ph/9704019.
- Vats, N., John, S. and Busch, K., 2002, Phys. Rev. A **65**, 043808.
- Verdú, S., 1998, IEEE Trans. Inf. Theory **44**, 2057.
- Verdú, S., and Han, T. S., 1994, IEEE Trans. Inf. Th. **40**, 1147.
- Weedbrook, C., Pirandola, S., Garcia-Patron, R., Cerf, N. J., Ralph, T. C., Shapiro, J. H. and Lloyd, S., 2012, Rev. Mod. Phys. **84**, 621.
- Wilde, M. M. and Brun, T. A., 2009, Phys. Rev. A **79**, 032313.
- Wilde, M. M. and Brun, T. A., 2010, Phys. Rev. A **81**, 042333.
- Wilde, M. M. and Brun, T. A., 2010, Quantum Inf. Processing **9**, 509.
- Wilde, M. M., Houshmand, M. and Hosseini-Khayat, S., 2011, Proceedings of the International Symposium on Information Theory, pp. 450-454.
- Winter, A., 1999, IEEE Trans. Info. Theory **45**, 2481.
- Wolf, M. M., Ortiz, G., Verstraete, F., and Cirac, J. I., 2006, Phys. Rev. Lett. **97**, 110403.
- Wolf, M. M., Pérez-García, D. and Giedke, G., 2007, Phys. Rev. Lett. **98**, 130501.
- Wolf, M. M. and Pérez-García, D., 2007, Phys. Rev. A **75**, 012303.
- Wolf, M. M. and Cirac, J. I., 2008, Commun. Math. Phys. **279**, 147.
- Wolfowitz, J., 1957, Illinois J. Math. **1**, 591.
- Ziman, M. and Buzek, V., 2005, Phys. Rev. A **72**, 022110.
- Zwanzig, R., 1960, J. Chem. Phys. **33**, 1338.